Université de Montréal

Segmentation des images radiographiques à rayon-X basée sur la fusion entropique et Reconstruction 3D biplanaire des os basée sur la modélisation statistique non-linéaire

par

Dac Cong Tai Nguyen

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Thèse présentée en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Discipline

August 27, 2022

Université de Montréal

Faculté des arts et des sciences

Cette thèse intitulée

Segmentation des images radiographiques à rayon-X basée sur la fusion entropique et Reconstruction 3D biplanaire des os basée sur la modélisation statistique non-linéaire

présentée par

Dac Cong Tai Nguyen

a été évaluée par un jury composé des personnes suivantes :

Jean Meunier
(président-rapporteur)

Max Mignotte
(directeur de recherche)

Frédéric Lavoie
(codirecteur)

Sébastien Roy
(membre du jury)

Luc Duong
(examinateur externe)

Jean Meunier
(représentant du doyen de la FESP)

Résumé

Dans cette thèse, nous présentons une méthode de segmentation d'images radiographiques des membres inférieurs en régions d'intérêt (ROIs), une méthode de recalage rigide tridimensionnel (3D) / bidimensionnel (2D) des prothèses du genou sur les deux images biplanaires radiographiques calibrées et une méthode de reconstruction 3D des membres inférieurs à partir de deux images biplanaires radiographiques calibrées.

Le premier article présente une méthode de segmentation de rotule, astragale et bassin des images radiographiques en régions d'intérêt basée sur la fusion de multi-atlas et superpixels. Cette méthode utilise l'apprentissage d'une base de données d'images radiographiques de ces os segmentées manuellement et recalées entre elles pour estimer un ensemble de superpixels permettant de tenir compte de toute la variabilité locale et non linéaire existante dans la base, puis la propagation d'étiquettes basée sur le concept d'entropie pour raffiner la carte de segmentations en régions internes afin d'obtenir le résultat final.

Le deuxième article présente une méthode de recalage rigide 3D / 2D des composants tibiaux et fémoraux de prothèse du genou sur deux images biplanaires radiographiques calibrées. Cette méthode utilise une mesure de similarité hybride basée sur les notions de contours et régions puis un algorithme d'optimisation stochastique pour estimer la position des composants. La similarité basée sur les régions est stable et robuste contre les bruits. Cependant, cette mesure n'est pas précise car le nombre de pixels aux contours est inférieur au celui à l'intérieur de la région. Au contraire, la similarité basée sur les contours est précise mais plus sensible au bruit ou à d'autres artefacts existant dans les images. C'est pourquoi la combinaison de ces deux similarités fournit une méthode de recalage robuste et précise.

Le troisième article représente une méthode statistique biplanaire de reconstruction 3D de rotule, astragale et bassin. Cette méthode utilise un algorithme de réduction de dimensionnalité pour définir un modèle déformable paramétrique qui contient toutes les déformations statistiques admissibles apprises à partir d'une base de données des structures osseuses. Puis un algorithme d'optimisation stochastique est utilisé pour minimiser la différence entre la projection des contours / régions des modèles surfaciques osseux avec ceux segmentés sur les deux images radiographiques.

Mots clés : Structures osseuses du membre inférieur, segmentation osseuse, segmentation consensus, segmentation multi-atlas, carte superpixel, étape de fusion basée sur la variation d'information, recalage 3D/2D, images radiographiques, composants d'implants de genou, implants orthopédiques, reconstruction 3D, radiographies biplanaires, modèles statistiques non linéaires, imagerie médicale.

Abstract

In this thesis, we present a segmentation method of lower limbs of X-ray images into regions of interest (ROIs), a three-dimensional (3D) / two-dimensional (2D) rigid registration method of knee implant components to biplanar X-ray images, and a 3D reconstruction method of the lower limbs using biplanar X-ray images.

The first paper presents a superpixel and multi-atlas-based segmentation method of the patella, talus, and pelvis into regions of interest. This method uses a training dataset of pre-segmented and co-registered X-ray images of these bones to estimate a collection of superpixels allowing to take into account all the nonlinear and local variability existing in the dataset, then a propagation of label based on the entropy concept for refining the segmentation map into internal regions to the final result.

The second paper presents a 3D / 2D rigid registration method of tibial and femoral components of knee implants to calibrated biplanar X-ray images. This method uses a hybrid edge- and region-based similarity measure then a stochastic optimization algorithm to estimate the component position. The region-based similarity is stable and robust to noise. However, this measure is not precise because the number of pixels in the border is fewer than the number of pixels inside the region. On the contrary, the edge-based similarity is accurate but more sensitive to noise or other artifacts existing in the images. That's why the combination of these two similarity types provides a robust and accurate registration method.

The third paper presents a statistical biplanar 3D reconstruction method of the patella, talus, and pelvis. This method uses a dimensionality reduction algorithm to define a deformable parametric model which contains all admissible statistical deformations learned from the bone structure dataset. Then a stochastic optimization algorithm is used to minimize the difference between the contour / region projection of bone models and the contours / regions in two segmented X-ray images.

Keywords : Bone structures of the lower limb, bone segmentation, consensus segmentation, multi-atlas segmentation, superpixel map, variation of information based fusion step, X-ray images, 3D/2D registration, knee implant components, orthopaedic implants, 3D reconstruction, biplanar radiographies, nonlinear statistical models, medical imaging.

Table des matières

Liste des tableaux

Liste des figures

Liste des sigles et des abréviations

2D	Bidimensionnel, de l'anglais *Two-dimensional*
3D	Tridimensionnel, de l'anglais *Three-dimensional*
CT-Scan	Tomodensitométrie, de l'anglais *Computed tomography scan*
MRI	Imagerie par résonance magnétique, de l'anglais *Magnetic resonance imaging*
CCD	Charged-coupled device
ML	Maximum de vraisemblance, de l'anglais *Maximum likelihood*
EM	Espérance maximisation, de l'anglais *Expectation maximization*
DOF	Degré de liberté, de l'anglais *Degree of freedom*
TKA	Arthroplastie totale du genou, de l'anglais *Total knee arthroplasty*
RSA	Roentgen stereophotogrammetric analysis
SLIC	Simple linear iterative clustering
NCC	Coefficient de corrélation normalisé, de l'anglais *Normalized correlation coefficient*
DLT	Transformation linéaire directe, de l'anglais *Direct linear transformation*

SSM	Modèle statistique de formes, de l'anglais *Statistical shape model*
SDM	Modèle statistique de déformation, de l'anglais *Statistical deformable model*
ES	Exploration sélection, de l'anglais *Exploration selection*
ICP	Iterative closest point
PA	Postéro-antérieur, de l'anglais *Postero-anterior*
LAT	Latéral, de l'anglais *Lateral*
CNN	Réseau de neurones convolutif, de l'anglais *Convolutional neural network*
GAN	Réseau antagoniste génératif, de l'anglais *Generative adversarial network*
DRR	Radiographie reconstruite numériquement, de l'anglais *Digitally reconstructed radiography*
PCA	Analyse en composantes principales, de l'anglais *Principal component analysis*
PPCA	Analyses en composantes principales probabilistes, de l'anglais *Probabilistic principal component analysis*
MDS	Positionnement multidimensionnel, de l'anglais *Multidimensional scaling*
LLE	Locally linear embedding
SDE	Semidefinite embedding

t-SNE	t-Distributed stochastic neighbor embedding
CPD	Coherent point drift
LSQ	Moindres carrés, de l'anglais *least squares*
ROI	Région d'intérêt, de l'anglais *Region of interest*
RMS	Moyenne quadratique, de l'anglais *Root mean square*
RMSE	Erreur quadratique moyenne, de l'anglais *Root mean square error*
AE	Erreur moyenne, de l'anglais *Average error*
SD	Écart type, de l'anglais *Standard deviation*
PB	Probabilistic patch-based label fusion model
MAPM	Multi-atlas Patch-Match
MASP	Multi-atlas spectral Patch-Match
DCT	Transformée en cosinus discrète, de l'anglais *Discrete cosine transform*
MV	Patch-based majority voting fusion scheme
VoI	Variation de l'information, de l'anglais *Variation of information*

Remerciements

Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.

Je voudrais dans un premier temps remercier grandement mon directeur de thèse Max Mignotte, professeur au Département d'Informatique et de Recherche Opérationnelle (Diro) de l'Université de Montréal, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter ma réflexion.

Je remercie docteur Frédéric Lavoie, chirurgien orthopédiste au Centre Hospitalier de l'Université de Montréal (CHUM) et président de la compagnie Eiffel Medtech Inc., qui m'a accepté comme stagiaire au doctorat au sein de sa compagnie. Il m'a ainsi offert des conseils et une expérience capitale pour mon avenir professionnel.

Je remercie monsieur Said Benameur, directeur de recherche et développement à Eiffel Medtech Inc., pour sa disponibilité, ses encouragements, son regard critique et sa rigueur scientifique. C'est à ses côtés que j'ai compris ce que rigueur et précision voulaient dire.

Je désire aussi remercier les professeurs du Département d'Informatique et de Recherche Opérationnelle (Diro) de l'Université de Montréal qui m'ont fourni les outils théoriques nécessaires à la réussite de mes études universitaires.

Je tiens à exprimer toute ma gratitude aux membres du jury d'avoir accepté mon invitation à juger ce travail.

Je tiens à remercier également toutes les personnes que j'ai pu côtoyer au sein de l'entreprise Eiffel Medtech et du laboratoire de traitement d'images pour l'aide qu'elles m'ont toujours prodiguée.

Je remercie ma chère famille pour ses encouragements, et qui a toujours été là pour moi.

Introduction

Contexte

La rotule, l'astragale et le bassin jouent un rôle crucial dans la physiologie et le fonctionnement des articulations du membre inférieur. La rotule est un petit os plat, triangulaire situé à la partie antérieure du genou. Elle fait partie de l'articulation et du mécanisme extenseur du genou. Elle protège l'avant de l'articulation du genou et permet de stabiliser le genou lors des différents mouvements de flexion et d'extension. L'astragale est un os court du pied en forme de poulie situé au-dessous du tibia. Il est le centre de l'articulation de la cheville et la liaison entre celles de la cheville et du pied. Il est important pour le squelette parce qu'il sert à soutenir le poids du corps humain, notamment durant la marche. Il est aussi important dans les mouvements latéraux et de flexion-extension du pied. Le bassin est une ceinture osseuse située sous l'abdomen qui soutient la colonne vertébrale. Il supporte le poids du haut du corps et protège les organes génitaux internes, la vessie, etc. Les os du bassin s'articulent avec ceux de la cuisse pour permettre la marche.

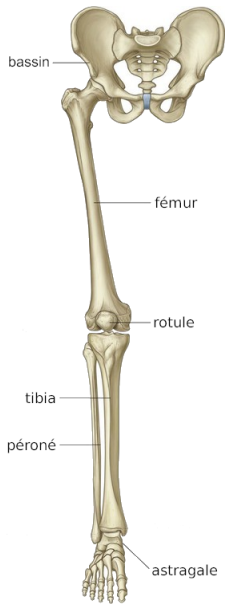


Fig. 1. Le membre inférieur.

Ces os entrent dans les activités quotidiennes de l'être humain durant toute sa vie. C'est pourquoi ils risquent de subir des pathologies (fractures, arthroses ...) plus fréquentes. Le diagnostic de ces pathologies nécessite le plus souvent de l'imagerie médicale. L'imagerie médicale se réfère à différentes technologies de création des représentations visuelles du corps humain. Chaque type de technologie fournit des informations différentes sur la partie du corps étudiée ou traitée. Elle est très utile pour les études cliniques et biomécaniques, les diagnostics et les plans de traitement. La reconstruction 3D de l'anatomie osseuse à l'aide d'images médicales, une application complexe de l'imagerie médicale, est utilisée depuis très longtemps et de plus en plus requise dans l'analyse clinique ainsi que dans le diagnostic et les traitements pré et postopératoires. Ainsi, une forme osseuse 3D précise peut améliorer la planification préopératoire dans les systèmes de navigation chirurgicale [**87**], ou augmenter la précision et la fiabilité en chirurgie orthopédique [**149**].

Problématique

Nombreuses méthodes de reconstruction 3D ont été proposées au cours des deux dernières décennies. La tomodensitométrie (CT-Scan) [**151, 86, 144**] et l'imagerie par résonance magnétique (MRI) [**94, 92**] permettent aux méthodes de reconstruction 3D de fournir un modèle 3D personnalisé précis. Les modèles générés à partir le CT-Scan sont généralement considérés comme un modèle de vérité terrain.

Cependant, le CT-Scan implique une exposition significative de radiation aux patients. La MRI n'est pas irradiante car elle n'utilise pas de rayons X ni de rayonnement ionisant. La MRI utilise des champs magnétiques puissants, des gradients de champs magnétiques et des ondes radio pour créer des images, alors les patients portant des implants médicaux ou de métal non amovibles peuvent ne pas subir un examen MRI en toute sécurité. De plus, les deux technologies (CT-Scan et MRI) sont coûteuses, produisent un volume important d'information à traiter et nécessitent des patients en position allongée, ce qui est incompatible avec certains protocoles de diagnostic. Si une série de suivis sont nécessaires, le coût et le problème d'irradiation dans le cas du CT-Scan augmenteront considérablement après une séquence d'examen. C'est pourquoi ils ne sont pas très pratiques. Une autre technique d'imagerie, complètement inoffensive pour le patient, est l'imagerie par ultrasons [**157**]. Cette dernière technique, non-invasive, n'a cependant pas été adaptée à l'analyse de la forme géométrique 3D des os car celles-ci génèrent trop d'écho et beaucoup de bruit de mesures (échos parasites à la forme géométrique des surfaces des structures denses).

Au contraire, la radiographie, dite «conventionnelle» utilisant les rayons X, est un choix intéressant car elle est la plus faible en termes de coût, de complexité et de risque d'irradiation par rapport à la MRI ou au CT-Scan (voir Tableau 1 [**53**]). Cependant, en raison de la variété du bruit d'imagerie provenant de plusieurs composants du système (source de rayons

X, caméra CCD, etc.) et du chevauchement entre les os et les tissus mous denses (cartilage, ménisque et fascia) du patient, les méthodes de reconstruction 3D multiplanaires restent un problème difficile.

	MRI	**CT-Scan**	**Radiographie (rayon X)**
Dimensionnalité	3D	3D	2D
Nombre d'images	≥ 400	≥ 700	≥ 2
Cartilage	oui	non	non
Rayonnement ionisant	non	3-8 mSv	0.04 mSv
Coût (moyenne nationale des États-Unis)	\$ 2500	\$ 1100	\$ 300

Tableau 1. Comparaison de la reconstruction du genou utilisant différentes modalités d'imagerie.

Le problème de reconstruction 3D multiplanaire, qui consiste à reconstruire en 3D un objet anatomique à partir de plusieurs vues radiographiques, appartient à la classe des problèmes mal posés au sens de Hadamard. Par exemple, l'acquisition de deux images radiographiques biplanaires ne fournit qu'un ensemble de données incomplètes qui ne permet pas d'assurer l'unicité de la solution. Pour pallier cette limitation, il est nécessaire de régulariser la procédure de reconstruction 3D en introduisant une connaissance *a priori* de l'objet à reconstruire. Des modèles géométriques ont été proposés mais leur grande rigidité ne permet pas de prendre en compte leur variabilité anatomique intrinsèque et celle liée à leurs pathologies [**44**] [**10**]. Au contraire, les méthodes de reconstruction 3D multiplanaires par modélisation statistique, telles que celle proposées par Benameur *et al.* dans [**17**] [**18**], offrent une plus grande souplesse de modélisation et permettent de prendre en compte les variabilités anatomiques et celles liées aux pathologies des structures osseuses à reconstruire.

La méthode développée par Benameur *et al.* dans [**17**] utilisait des contours détectés sur les deux images radiographiques biplanaires et une connaissance géométrique *a priori* de nature statistique de l'objet à reconstruire obtenue par une analyse en composantes principales (PCA) et/ou un mélange d'analyse en composantes principales probabiliste (PPCA) appliquée à une base d'apprentissage. Cette dernière consistait à ajuster les projections (postéro-antérieure et latérale) de ce modèle 3D de l'objet à reconstruire avec les contours préalablement segmentés de l'objet correspondante sur les deux images radiographiques. Le problème de reconstruction 3D était considéré comme un problème de minimisation d'une fonction de coût optimisée par l'algorithme Exploration Sélection (ES).

Les méthodes linéaires ne sont pas efficaces lorsque les données ne sont pas distribuées dans un sous-espace linéaire de l'espace original. Pour pallier ce problème, de nombreux algorithmes de réduction de dimensionnalité non linéaires ont été développées [**69**] [**168**] [**155**]. Ces algorithmes sont généralement basés sur l'idée qu'il existe un sous-espace non linéaire, qui contient la distribution des données. Parmi les algorithmes proposés dans la littérature, les plus populaires sont l'Isomap [**155**], le positionnement multidimensionnel

(MDS) [**26**], Locally Linear Embedding (LLE) [**137**], Semidefinite Embedding (SDE) [**95**] et t-Distributed Stochastic Neighbor Embedding (t-SNE) [**169**]. Dans ce projet, nous nous intéresserons à la méthode FastMap qui est une variante de MDS. Il existe des variantes non métriques de la MDS, qui conservent les relations ordinales dans les données, au lieu des paires de distances [**41**]. Parmi les avantages de la méthode MDS par rapport aux autres méthodes de réduction de dimensionnalité non linéaires est sa capacité d'être très flexible dans le choix de la mesure de distance.

Dans les images radiographiques, l'extraction des contours des structures osseuses (rotule, astragale, et bassin) n'est pas une tâche facile à cause du bruit élevé, du faible contraste et de la superposition de structures. Les algorithmes de détection de contours sont très sensibles au bruit et il en résulte souvent des contours bruités, difficilement exploitables dans le cas des images issues du système de radiographie biplane basse-dose qui seront utilisées dans notre projet.

Une approche plus robuste consistera en une méthode de segmentation qui permet de prendre en compte à la fois la notion de contour et de région (détection de zones homogènes). La coopération entre les segmentations par régions et par contours contribue à une meilleure prise en compte des caractéristiques des os dans les images radiographiques et, par conséquent, à une meilleure segmentation [**23**] grâce à la nature complémentaire de ces deux types d'information [**177**]. Ainsi, une segmentation par coopération régions-contours peut être exprimée comme une entraide entre ces deux concepts afin d'améliorer le résultat final de segmentation [**42**] [**120**].

À cette fin, nous proposerons une méthode de segmentation d'images basse-dose, basée sur la fusion entropique en utilisant les superpixels et multi-atlas.

Objectifs

Pour pallier ces limitations, les objectifs de cette thèse sont de développer une méthode de segmentation des membres inférieurs des images radiographiques, une méthode de recalage rigide 3D / 2D des prothèses du genou sur les deux images biplanaires, ainsi une méthode de reconstruction 3D des membres inférieurs à partir de deux images biplanaires. Les images utilisées dans ces méthodes sont les images radiographiques à rayon X calibrées. Les membres inférieurs considérés dans cette thèse sont la rotule, l'astragale et le bassin. Ces os ont un rôle important, mais il y a peu d'études dans la littérature sur eux comparativement aux os longs (fémur, tibia et péroné).

À cette fin, nous avons développé une méthode de segmentation des os (rotule, astragale et bassin) des images radiographiques en régions basées sur la fusion de multi-atlas et superpixels, une méthode de recalage rigide 3D / 2D des composants tibiaux et fémoraux de la prothèse du genou sur les deux images biplanaires radiographiques calibrées et une méthode

de reconstruction 3D des rotules, astragales et bassins à partir de deux images biplanaires radiographiques calibrées. La reconstruction 3D utilise la première méthode pour segmenter les images en entrées afin d'extraire les régions osseuses des images, et la deuxième méthode pour positionner au préalable le modèle paramétrique dans l'environnement radiographique. Ensuite, un algorithme d'optimisation est appliqué pour minimiser une fonction d'énergie afin d'estimer des paramètres de déformation du modèle (voir Fig. 2).

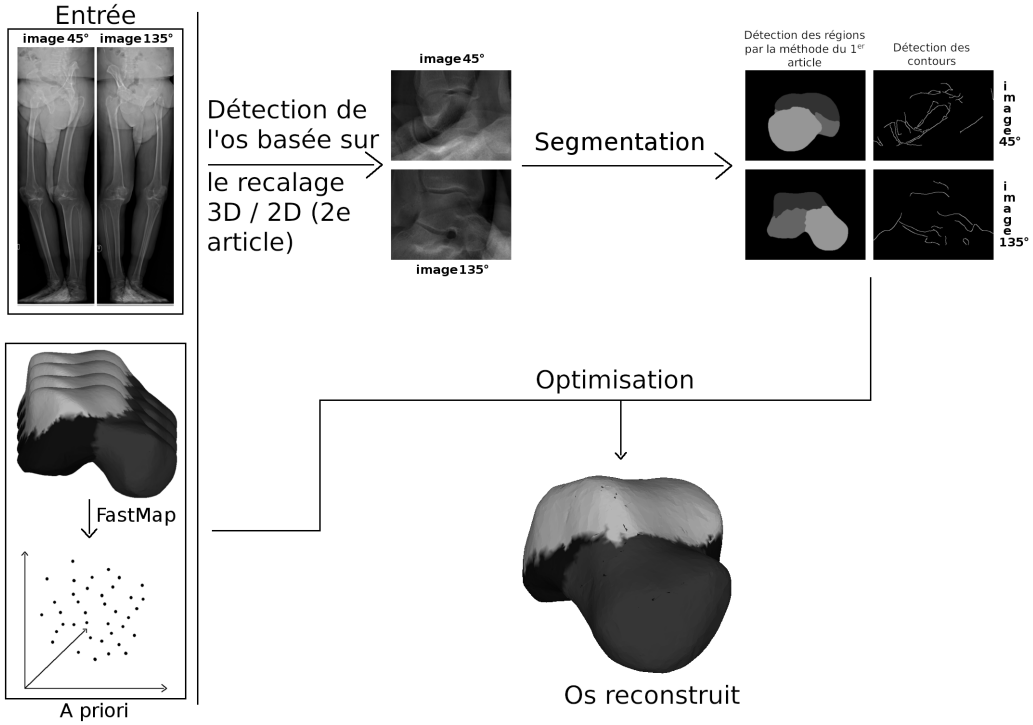


Fig. 2. Le schéma de la thèse.

Le premier article présente une méthode de segmentation de rotule, astragale et bassin des images radiographiques en régions d'intérêt basée sur la fusion de multi-atlas et superpixels. Cette méthode utilise l'apprentissage d'une base de données des images radiographiques de ces os segmentées manuellement et recalées entre elles pour estimer un ensemble de superpixels permettant de tenir compte de toute la variabilité locale et non linéaire existant dans la base. Un processus d'élagage est ensuite appliqué sur cet ensemble afin de définir les contours externes et une propagation d'étiquettes basée sur le concept d'entropie pour raffiner la carte de segmentations en régions internes afin d'obtenir le résultat final.

Le deuxième article présente une méthode de recalage rigide 3D / 2D des composants tibiaux et fémoraux de la prothèse du genou sur les deux images biplanaires radiographiques calibrées. Cette méthode utilise une mesure de similarité hybride basée sur les notions de contours et régions puis un algorithme d'optimisation stochastique pour estimer la position des composants. La similarité basée sur les régions est stable et robuste contre les bruits.

Cependant, cette mesure n'est pas précise car le nombre de pixels aux contours est inférieur à celui à l'intérieur de la région. Au contraire, la similarité basée sur les contours est précise mais plus sensible au bruit ou à d'autres artefacts existant dans les images. C'est pourquoi la combinaison de ces deux similarités fournit une méthode de recalage robuste et précise.

Le troisième article représente une méthode statistique biplanaire de reconstruction 3D de rotule, astragale et bassin. Cette méthode utilise un algorithme de réduction de dimensionnalité pour définir un modèle déformable paramétrique qui contient toutes les déformations statistiques admissibles apprises à partir d'une base de données des structures osseuses. Puis un algorithme d'optimisation stochastique est utilisé pour minimiser la différence entre la projection des contours / régions des modèles surfaciques osseuses avec ceux segmentés sur les deux images radiographiques.

Plan de la thèse

Le chapitre "Revue de littérature" décrit une recherche de littérature pour illustrer la contribution des travaux précédents à notre thèse.

Les trois chapitres suivants correspondent aux trois articles développés dans cette thèse:

"Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images", publié dans le journal Medical Image Analysis (MEDIA), Elsevier, vol. 48 : 58-74, August 2018

"Unsupervised Registration of 3D Knee Implant Components to Biplanar X-ray Images", soumis au journal BMC Medical Imaging, 2021

"3D Biplanar Reconstruction Of Lower Limb Using Nonlinear Statistical Models", soumis au journal Elsevier Medical Image Analysis, 2022

Nous discutons dans le chapitre "Discussion générale" et concluons dans le chapitre "Conclusion générale"

Revue de littérature

Comme il a été mentionné dans le chapitre précédent, notre thèse est constituée de trois parties: La segmentation des images radiographiques en ROIs, le recalage 3D/2D et la reconstruction 3D biplanaire des os des membres inférieurs par modélisation statistique non-linéaire. Dans ce chapitre, nous présentons une revue de littérature pour chacun de ces thèmes: segmentation, recalage et reconstruction 3D.

Segmentation

L'extraction des contours des structures osseuses n'est pas une tâche facile à cause du bruit élevé, du faible contraste et de la superposition de structures. De nombreuses techniques de segmentation des images radiographiques à rayon X ont été proposées dans la littérature. Parmi ces techniques, on peut citer la technique des contours actifs [**80**], Level Set [**145**], technique utilisant un modèle 3D générique [**61**], technique basée sur une combinaison de Level Set et Live-Wire [**14, 154**], technique combinant les contours actifs avec Level Set pour introduire de l'information *a priori* et des contraintes de forme pour déformer le contour initial [**54**], technique basée sur modèle statistique [**15**]. Chen *et al* [**37**] ont eu recours à un modèle *a priori* 2D pour extraire le contour du fémur proximal. Smith *et al* [**150**] ont proposé un modèle statistique 2D pour segmenter le bassin et le fémur proximal. Ding *et al* [**48**] ont proposé une méthode de segmentation du bassin et du fémur proximal en utilisant une approche basée sur un atlas combiné par une déformation par Level Set. Hill *et al* [**68**] ont proposé une méthode pour construire un gradient de texture à partir d'une transformation par ondelettes. Vanhamel *et al* ont introduit dans [**162**] une approche marginale pour appliquer des filtres de Gabor à chaque composante d'une image couleur et ainsi construire un espace de caractéristiques couleurs/textures utilisées dans la segmentation. De manière analogue, Hoang *et al* utilisent des filtres de Gabor pour mesurer la texture-couleur et la segmentation est obtenue avec une classification par k-moyens [**70**]. Les travaux de l'équipe Malik *et al* s'appuient aussi sur une base de filtres gaussiens pour calculer un gradient de texture qui est ensuite combiné avec des gradients de brillance et de couleur dans un schéma d'apprentissage supervisé [**98**].

Ces algorithmes de segmentation (détection de contours) sont très sensibles au bruit et il en résulte souvent des contours bruités, difficilement exploitables dans le cas des images issues du système de radiographie biplane basse-dose qui seront utilisées dans notre projet.

Une approche plus robuste consistera en une méthode de segmentation qui permet de prendre en compte à la fois la notion de contour et de région (zones homogènes). La combinaison des deux segmentations par régions et par contours nous permet une meilleure prise en compte des caractéristiques des os dans les images radiographiques et une meilleure segmentation [**23**] grâce à la nature complémentaire de ces deux types d'information [**177**]. Ainsi, une segmentation par coopération régions-contours peut être exprimée comme une entraide entre ces deux concepts afin d'améliorer le résultat final de segmentation [**42, 120**].

Recalage 3D/2D

Les méthodes de recalage 3D / 2D des implants sont classées suivant le nombre d'images utilisé. Le premier groupe utilise une seule image à rayon X et des images à rayon X biplanaires sont utilisées dans le second groupe.

Dans le premier groupe, Mahfouz *et al* [**96**] a utilisé une mesure de similarité d'image à image et un algorithme d'optimisation robuste pour éviter les minima locaux et trouver une bonne solution. Yamazaki *et al* [**173**] a utilisé un recalage 3D/2D basé sur les contours pour estimer la pose de l'implant du genou. Un recalage 3D/2D statistique basé sur un algorithme Maximum de vraisemblance (ML) - Espérance-maximisation (EM) a été proposé par Hermans *et al* [**67**]. Il aligne le contour du composant avec les contours extraits d'une image. De telles techniques ne peuvent fournir une précision clinique suffisante que pour cinq degrés de liberté (DOF) (trois rotations et deux translations parallèles à l'image), car le DOF lié à la translation perpendiculaire à l'image (profondeur) est assez difficile à estimer. Yamazaki *et al* [**173**] a amélioré l'estimation de la profondeur en l'optimisant indépendamment des cinq autres DOFs, en utilisant une courbe d'évaluation approximative de la profondeur préparée après le recalage initial. Bien que la profondeur ait été améliorée, elle n'a pas été suffisamment précise. Une autre approche pour augmenter la précision de l'estimation de la profondeur était basée sur la propriété d'imagerie fluoroscopique selon laquelle plus l'objet est proche de la source, plus l'image à l'intensificateur d'image est grande. Hossain *et al* [**71**] a déterminé le changement d'échelle dû à la translation de profondeur en utilisant une boîte de calibration pour estimer la profondeur. Cette approche a donné des résultats précis mais a nécessité une étape supplémentaire pour calculer le changement d'échelle dans la translation en profondeur de la fluoroscopie. La deuxième limitation est due à la symétrie de la forme des composants, en particulier le composant tibial. Une position symétrique peut être obtenue au lieu de la vraie car la projection des deux solutions sur l'image est très similaire.

Dans [**67**], les auteurs ont tenté de résoudre ce problème en estimant simultanément deux poses symétriques, mais l'algorithme peut toujours ne pas converger vers la vraie solution.

En utilisant des images biplanaires, ces méthodes peuvent mesurer les six DOFs avec une précision suffisante et éviter le problème de symétrie. Kim *et al* [**83**] a optimisé le coefficient de corrélation normalisé (NCC) entre les images à rayon X et deux images de projection des composantes correspondantes pour obtenir six DOFs. George *et al* ont montré que l'utilisation d'images biplanaires fournit une estimation très précise de six DOFs [**64**]. Cependant, les images fluoroscopiques biplanaires à la même position ne sont pas faciles à obtenir en raison du mouvement des images et de la complexité du système. De plus, la fluoroscopie implique un risque potentiel de radiation pour le patient.

Contrairement aux autres méthodes, notre méthode utilise des images radiographiques biplanaires qui sont avantageuses en termes de coût, de complexité et de risque de rayonnement et ne nécessitent pas de marqueur fiduciaire. Notre méthode est également robuste au bruit d'image et aux occlusions grâce à une mesure de similarité hybride basée sur les contours et les régions. Cette mesure permet d'améliorer la stabilisation et la précision des approches grâce à la robustesse contre les bruits et la stabilisation de la similarité basée sur les régions et la précision de la similarité basée sur les contours.

Reconstruction 3D

La reconstruction 3D commence par le groupe utilisant la notion de correspondance stéréo/non stéréo et la transformation linéaire directe (DLT), proposée par Abdel-Aziz et Karara [**1**] (voir détails dans [**99**]). La technique DLT utilise les coordonnées bidimensionnelles (2D) des points correspondants d'au moins deux images radiographiques différentes pour estimer leur position 3D. Dans [**126, 44, 6**], les auteurs ont identifié quelques points de repère anatomiques stéréo-correspondants, puis ont utilisé l'algorithme DLT pour générer des points 3D de la colonne vertébrale humaine et de la cage thoracique, respectivement. Ces méthodes ne fournissent pas un modèle 3D fiable et précis car le nombre de points de repère utilisés était trop faible et seules les informations locales ont été prises en compte. Notez qu'il n'est pas simple d'ajouter plus de points de repère en raison des difficultés à identifier les points de repère correspondants dans deux images.

Pour augmenter la précision, le deuxième groupe a introduit l'utilisation d'un objet générique déformé par un algorithme de krigeage pour obtenir la reconstruction 3D. Mitton *et al* [**116**] ont utilisé des points correspondants non stéréo et un objet générique déformable pour ajouter plus de points et considérer l'information globale de la forme des vertèbres cervicales. Cette méthode reconstruisait les points correspondants stéréo et non stéréo, puis les utilisait comme points de contrôle pour le krigeage de l'objet générique. L'approche a produit une reconstruction 3D plus précise que les méthodes susmentionnées. Cependant,

le nombre limite de points correspondants ne peut pas représenter correctement une forme continue. Pour cette raison, Laporte *et al* [**89**] ont proposé une approche de contour correspondant non stéréo au lieu de points de repère. Le principe de la méthode est le processus d'association des points de contrôle de la projection des contours des régions anatomiques de l'objet générique à des points de contrôle de contours 2D identifiés dans deux images, et la minimisation de la distance entre deux ensembles de points de contrôle. Les régions anatomiques ont été définies manuellement sur l'objet générique. Les contours 2D ont été identifiés manuellement/semi-manuellement sur deux radiographies. Ensuite, un algorithme de krigeage a été appliqué en considérant les points de contrôle pour obtenir le résultat final. Le résultat obtenu a montré que la méthode fournit une reconstruction 3D considérablement plus précise dans le cas du fémur distal [**89**], du fémur proximal [**91**], du bassin [**117**] et de la cage thoracique [**24**]. Pour améliorer la reconstruction, notamment l'initialisation, Chaibi *et al* [**35**] et Quijano *et al* [**129**] ont introduit un modèle morpho-réaliste paramétrique personnalisé en régionalisant et en paramétrant les principales caractéristiques de l'objet générique par des primitives géométriques telles que des points 3D, des sphères, des cylindres, des segments 3D, etc. Les limites de ce groupe sont la dépendance des compétences de l'expert pour identifier exactement des points de repère/contours, la sensibilité à la distinction des caractéristiques géométriques dans les radiographies, le nombre limite de points de correspondants/contrôle ne peut pas représenter correctement une forme continue, et l'objet générique a moins de détails que l'objet réel.

Ces limitations conduisent au troisième groupe qui utilise le modèle statistique de formes (SSM). Les méthodes basées sur SSM réduisent l'intervention de l'utilisateur et augmentent la précision en prenant en compte la structure géométrique globale des objets. Les connaissances géométriques a priori proviennent d'une large base de données d'apprentissage d'objets incluant des sujets normaux et pathologiques, et représentent toutes les déformations de formes possibles. Ces informations sont utilisées comme contraintes géométriques de la forme de l'objet. Le SSM a été utilisé avec succès en imagerie médicale [**38, 57, 93, 7, 43, 143, 79**]. Ainsi, Fleute et Lavallée [**58**] ont présenté un algorithme de reconstruction 3D du fémur distal utilisant une généralisation de l'algorithme Iterative Closest Point (ICP) pour recaler le modèle aux contours segmentés dans quelques radiographies. Benameur *et al* [**19**] ont proposé une méthode de reconstruction 3D des vertèbres du rachis scoliotique utilisant deux images radiographiques conventionnelles (postéro-antérieur et latéral). Cette méthode a utilisé une technique de descente de gradient pour minimiser une fonction de coût basée sur les contours potentiels pour chaque vertèbre individuelle. Et la reconstruction de la colonne vertébrale s'est faite vertèbre par vertèbre. La méthode de Baka *et al* [**12**] était basée sur [**19**] pour reconstruire le fémur distal. Cette approche diffère en ce que leur algorithme incluait une pondération d'orientation et optimisait l'erreur 3D du SSM au lieu d'une erreur 2D des

contours projetés. Benameur *et al* [**18**] ont amélioré leur méthode en utilisant hiérarchiquement deux SSM de la structure géométrique de l'ensemble du rachis et de chaque vertèbre. Le premier modèle a donné une reconstruction grossière de l'ensemble de la colonne vertébrale, et le second a fourni une reconstruction précise de chaque vertèbre. De nombreuses méthodes basées sur la SSM ont été proposées pour les reconstructions du bassin [**178**], de la cage thoracique [**51**], de la colonne vertébrale [**119**], etc. Le point commun de ces méthodes est la nécessité d'une initialisation manuelle/semi-manuelle dans une étape préliminaire. Aubert *et al* [**9**] ont utilisé de différents réseaux de neurones convolutifs (CNN) pour l'initialisation automatique, puis des points de repère prédits par CNN ont été utilisés pour effectuer une déformation locale comme étape de raffinement. À notre connaissance, toutes les méthodes basées sur la SSM utilisaient l'algorithme d'analyse en composantes principales / mélanges d'analyses en composantes principales probabilistes [**17**].

Le quatrième groupe est que le problème de reconstruction 3D se présente comme un recalage 3D/2D non rigide entre les contours/régions projetés de l'objet générique avec les contours/régions extraits des images d'entrée. Gamage *et al* [**62**] ont utilisé une nouvelle mesure de similarité, qui est une combinaison pondérée d'informations de topologie, d'orientation des contours, de courbure et de distance euclidienne. Cette mesure est robuste contre le bruit, la déformation, les valeurs aberrantes et les occlusions dans les images radiographiques. Le recalage a été amélioré en impliquant non seulement les contours externes, mais aussi quelques bords internes clés du fémur et peut être adapté à d'autres anatomies osseuses. Sun *et al* [**153**] ont adapté l'algorithme Coherent Point Drift (CPD) [**121**] pour effectuer le recalage 3D/2D non rigide du fémur. L'algorithme CPD était basé sur la probabilité d'ensembles de points. À noter que dans les deux études, les auteurs n'ont pas mentionné comment l'initialisation de la reconstruction a été effectuée. L'avantage des techniques basées sur le recalage est de fournir un meilleur contrôle sur la déformation locale de la forme de l'objet. Cependant, ce groupe est très sensible à l'initialisation et la segmentation / détection des contours / régions des images radiographiques qui restent une tâche difficile.

Le dernier groupe utilise l'apprentissage profond, plus précisément les réseaux de neurones convolutifs, pour créer un modèle 3D. Récemment, CNN est utilisé pour estimer / reconstruire en 3D les objets [**63, 160**] ou les os d'animaux [**66, 33**] à partir des images 2D. Mais peu d'études ont été faites pour reconstruire en 3D les os humains. Shiode *et al* [**147**] utilise le réseau intégré TL sur une base d'apprentissage synthétique de nombreux modèles 3D avec des radiographies reconstruites numériquement (DRR) correspondantes. Le T et L se réfèrent respectivement à la phase d'apprentissage et celle de test. Le réseau T est un auto-encodeur avec des couches de convolution et de déconvolution qui permet de générer les modèles 3D et le réseau L est celui de T sans l'encodeur qui permet de prédire le modèle 3D à partir d'une image. La méthode commence par appliquer le réseau antagoniste génératif (GAN) sur la segmentation de l'os de l'image radiographique réelle en entrée afin de déduire

la DRR. Ensuite, cette DRR est fournie au réseau TL pour produire le modèle 3D de l'os. Néanmoins, cette méthode ne s'applique qu'aux os de l'avant-bras distal et surtout présuppose que ceux-ci soient sains. Kasten *et al* [**81**] introduisent une approche d'agrandissement dimensionnel qui est la rétroprojection de chaque paire de lignes épipolaires correspondantes des deux images radiographiques biplanaires en entrée sur un plan épipolaire. Il en résulte un volume 3D qui contient toutes les informations à partir des deux images, tout en préservant les contraintes géométriques de deux vues. Cette représentation est combinée ensuite avec CNNs qui produit des modèles 3D des différents os. Cependant, cette approche reste une tâche difficile, en raison de la difficulté de représenter un agrandissement dimensionnel dans des paramètres multi-vues avec des couches différentiables standards. De plus, en raison de la nature des images radiographiques (bruits d'imagerie, artefacts, chevauchement entre les os et les tissus mous denses), la correspondance des points de surface entre plusieurs vues est extrêmement difficile. L'accès difficile aux données médicales constitue une limitation importante de cette catégorie de méthode de reconstruction 3D basée sur les CNNs. En effet, la base d'apprentissage synthétique utilisée par ces méthodes et construite à partir de ces données réelles souvent limités, s'avère généralement trop petite et ainsi trop peu représentative pour aboutir à de bons résultats.

Pour limiter les inconvénients, notre approche utilise deux images radiographiques biplanaires obliques et la méthode statistique non linéaire FastMap pour créer un SDM pour la reconstruction 3D. Le FastMap permet de fournir une bonne représentation des déformations statistiques admissibles dans la base de données osseuses. Et les images obliques permettent à notre robuste fonction de vraisemblance d'exploiter toutes les informations des deux images (contours et ROIs).

Premier article.

Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images

par

Dac Cong Tai Nguyen[1], Said Benameur[2], Max Mignotte[3] et Frédéric Lavoie[4]

([1]) DIRO, Université de Montréal et Eiffel Medtech Inc
([2]) Eiffel Medtech Inc
([3]) DIRO, Université de Montréal
([4]) Département de Chirurgie Orthopédique, CHUM et Eiffel Medtech Inc

Cet article a été soumis dans Elsevier Medical Image Analysis.

Mes contributions et le rôle des coauteurs

- DCTN a réalisé les travaux et rédigé le manuscrit.
- SB, MM et FL ont supervisé techniquement et biomédicalement ces travaux, et révisé le manuscrit.

Résumé. La segmentation des images radiographiques est une étape importante et cruciale pour la reconstruction 3D osseuse dont l'objectif final reste d'augmenter l'efficacité du diagnostic, de la chirurgie et des plans de traitement assistés par ordinateur. Cependant, cette tâche de segmentation est plutôt difficile, en particulier lorsqu'il s'agit des structures humaines complexes du membre inférieur telles que la rotule, le talus et le bassin. Dans ce travail, nous présentons une méthode basée sur la fusion multi-atlas pour la segmentation automatique de ces régions osseuses complexes à partir d'une seule vue radiographique. La première originalité de l'approche proposée réside dans l'utilisation d'une base de données d'images radiographiques co-recalées/pré-segmentées des régions osseuses (ou multi-atlas) pour estimer une collection de superpixels permettant de prendre en compte toute la variabilité non linéaire et locale des régions osseuses existant dans la base de données et également pour simplifier le processus d'élagage de la carte superpixel lié à notre stratégie. La deuxième originalité est d'introduire une nouvelle étape de propagation d'étiquettes basée sur le concept d'entropie pour raffiner la carte de segmentation résultante en les régions internes les plus probables jusqu'à la segmentation finale. Dans le cadre de ce travail, une validation croisée *leave-one-out* a été effectué sur un ensemble de données de 31 images radiographiques segmentées manuellement pour chaque structure osseuse afin d'évaluer rigoureusement l'efficacité de la méthode proposée. La méthode proposée a abouti à des segmentations plus précises par rapport au *probabilistic patch-based label fusion model* (PB) et au *classical patch-based majority voting fusion scheme* (MV) utilisant différentes stratégies de recalage. La comparaison avec les segmentations manuelles (étalon-or) a révélé que la précision de classification de notre approche de segmentation non supervisée est, respectivement, de 93,79 % pour la rotule, de 88,30 % pour le talus et de 85,02 % pour le bassin ; un score qui se situe dans l'intervalle des niveaux de précision des segmentations manuelles (du fait de la variabilité intra inter/observateur).
Mots clés : Structures osseuses du membre inférieur, segmentation osseuse, segmentation consensus, segmentation multi-atlas, carte superpixel, étape de fusion basée sur la variation d'information, images radiographiques.

Abstract. X-ray image segmentation is an important and crucial step for three-dimensional (3D) bone reconstruction whose final goal remains to increase effectiveness of computer-aided diagnosis, surgery and treatment plannings. However, this segmentation task is rather challenging, particularly when dealing with complicated human structures in the lower limb such as the patella, talus and pelvis. In this work, we present a multi-atlas fusion framework for the automatic segmentation of these complex bone regions from a single X-ray view. The first originality of the proposed approach lies in the use of a (training) dataset of co-registered/pre-segmented X-ray images of these aforementioned bone regions (or multi-atlas) to estimate a collection of superpixels allowing us to take into account all the nonlinear and local variability of bone regions existing in the training dataset and also to simplify the superpixel map pruning process related to our strategy. The second originality is to introduce a novel label propagation step based on the entropy concept for refining the resulting segmentation map into the most likely internal regions to the final consensus segmentation. In this framework, a leave-one-out cross-validation process was performed on 31 manually segmented radiographic image dataset for each bone structure in order to rigorously evaluate the efficiency of the proposed method. The proposed method resulted in more accurate segmentations compared to the probabilistic patch-based label fusion model (PB) and the classical patch-based majority voting fusion scheme (MV) using different registration strategies. Comparison with manual (gold standard) segmentations revealed that the good classification accuracy of our unsupervised segmentation scheme is, respectively, 93.79% for the patella, 88.30% for the talus and 85.02% for the pelvis; a score that falls within the range of accuracy levels of manual segmentations (due to the intra inter/observer variability).
Keywords: Bone structures of the lower limb, bone segmentation, consensus segmentation, multi-atlas segmentation, superpixel map, variation of information based fusion step, X-ray images.

1. Introduction

X-ray images are used by physicians all over the world for the preliminary diagnosis of several bone diseases, to plan surgical intervention, and for pre and post-operative treatments. In this context, accurate extraction of bone contours or regions from these 2D X-ray images, is often the preliminary, and also crucial, step for three-dimensional (3D) bone reconstruction which can then be a great help in determining the extent of a fracture or for improving the diagnosis, follow-up, and treatment of major bone diseases, such as osteoporosis and osteoarthritis.

X-ray segmentation of bone structures of the lower limb, especially of the pelvic region, remains a challenging problem due both to intrinsic and extrinsic difficulties. Intrinsic difficulties refer to the intrinsic properties of the X-ray imaging systems that result in imaging noise. The major source of noise that degrades image quality is from the X-ray source and is caused by radiation scattering and source leakage (quantum noise). Extrinsic difficulties are closely related to the patients. Indeed, bone structures in radiographic images often overlap

with other bones or dense soft tissues, making it difficult to accurately delimit the boundaries of different bone regions. This is particularly true when neighboring structures have similar absorption rates. Also, we must include the effects of the density variation of the bone structure being imaged, the inter-subject variability in bone morphology, the variability in the 3D imaging poses, and the motion unsharpness (caused by movement of the patient). These problems are all the more difficult when dealing with large and complex bone shapes such as the human pelvis or small bone structures such as the talus or patella. In these cases, the segmentation of complex bones is usually still done manually or semi-automatically, which is time consuming and tedious.

In order to propose a fully automatic segmentation method of bone structures in X-ray images, some *a priori* anatomical knowledge of the region to be extracted is necessary and must be integrated in the segmentation model. This can be done in different ways, for example by modelling the grey level distributions associated with each tissue or region class to be segmented [**111**], in the form of deformable shape templates constrained by a family of parametric or non-parametric curves [**114, 112**], or by one or several prototype templates together with the set of admissible deformations [**45**].

Recently, an efficient example-based segmentation strategy has been proposed, which turns out to be an interesting alternative in order to incorporate prior information into the segmentation model. This alternative uses a training dataset of X-ray images with the corresponding pre-segmentations called a multi-atlas [**8, 5, 49**]. In this strategy, the unsupervised segmentation process is divided into two stages; a first registration procedure is conducted between the target image to be segmented and each image and/or segmentation of the multi-atlas. Next, from the set of co-registered segmentations (and possibly its corresponding X-ray image), a second *fusion* or *label propagation* step is finally achieved to estimate both the segmentation into regions and also to infer the most likely internal region labels to the final consensus segmentation result. Regarding the second stage of any multi-atlas based segmentation scheme, also called *label propagation* or *decision fusion* step, the commonly used techniques, already proposed in this context, are generally based on different types of voting rules, such as the majority voting [**133**], (possibly locally) weighted voting [**8**] in which weights can be possibly estimated by the Expectation-Maximization (EM) algorithm [**135**], or shape-based methods [**134**] to name a few.

The limits of these techniques are mainly due to the registration errors. Regions from atlases may be associated to the wrong regions in the target image. To address this limitation, a first strategy consists in using a combination of a linear rigid registration, followed by a nonlinear or local registration. The first step provides an initial rigid alignment, while the second step rather considers the specific non-rigid and/or nonlinear deformations of the target (for example, to take into account the inter-individual variability of the imaging pose and bone morphology). In this context, the nonlinear registration can be achieved in many ways.

For example, by considering small perturbations (as parameters of the nonlinear transform) on landmark points of a preliminary rigid registration, or by considering nonlinear deformations of a grid of Bezier or BSpline points, using deformation field (optical flows), demons, finite element models, to name a few. [**118**]. In addition to the fact that this nonlinear registration is often slow and suboptimal due to non-convexity of the energy function that is optimized, let us recall, above all, that the nonlinear deformation required for this registration is generally not learned from the multi-atlas dataset. A second strategy focuses on patch-based label fusion methods to compensate registration error by searching for correspondences between the target image and atlases. In this spirit, [**39**] use expert segmentation priors to achieve the nonlocal patch-based label fusion. [**136**] propose a patch-based image labeling method, relying on image intensity similarities between the input image and an anatomy textbook. [**166**] propose a regression-based approach for label fusion. Patch-based approach in [**170**] uses $l_{2,1}$-norm regularization to enforce joint sparsity during the label fusion. [**165**] combine (possibly weighted) majority voting and a patch-based approach to achieve unsupervised segmentation. [**11**] formulate patch-based label fusion in a probabilistic Bayesian framework. [**146**] combine spectral matching with multi-atlas Patch-Match (MAPM) to introduce multi-atlas spectral Patch-Match (MASP) segmentation. Finally, [**171**] propose a hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. Multi-scale feature representation increases the accuracy of the patch-based similarity measurement and label-specific atlas patches makes the label fusion process more specific and flexible. The interest of these methods is that they don't need explicit registration between the individual putative segmentations of the atlas and the subject image to be segmented. In fact, the main assumption of all these patch-based approaches is that, if patches of the input subject image are locally similar to the patches of atlases, they should have a similar label. In terms of voting rule, the atlases whose reference images are more similar to the target image should contribute more to the segmentation. This strategy borrows the main idea of the non-local means denoising algorithm [**27**] which assumes that image information is redundant.

In our approach, we prefer not to rely too much on patch-based similarity between two X-ray images mainly because grey level information of these patches are not always as informative as might be desired. Indeed, quantum noise can be different for different X-ray imaging systems and numerous artifacts exist, such as bone and soft-tissue artifacts and inter-individual variability of bone densities and structures. In fact, the most informative and reliable visual cues in an X-ray image of the pelvic region remain the boundary contours between the different pelvic bone structures, and particularly the external bone contour of the pelvis; our proposed method will fully rely on this information.

In our model, this boundary information between bone regions existing in the pre-segmented images of the multi-atlas dataset allows us to estimate a collection of superpixels

which has the interesting property to capture, easily and non-parametrically, all the nonlinear and local variability of Regions Of Interest (ROIs) present in the multi-atlas dataset. The interest is then twofold; this allows to efficiently estimate the target-specific nonlinear deformations required by the registration process while simplifying the optimization problem involved in the pruning process (*via* a superpixel map based refining process aiming at reducing the remaining registration errors). Then, the *label propagation* step is formulated into an optimization problem, also called the median partition problem, aiming to find the best segmentation which maximizes an entropy-based similarity criterion between the compromise solution and the set of co-registered segmentations, over the space of possible segmentations. This fusion procedure has already turned out to be very efficient for combining a set of weak segmentation results to get a final improved segmentation result [**109**].

Let us mention that the concept of superpixels was initially proposed in multi-atlas segmentation by [**167**] and [**175**] but with a very different approach compared to us, both in terms of data and algorithm used to generate the superpixel map but also in terms of their utility in the underlying registration/segmentation model. Indeed, in the above mentioned works, the superpixel representation is generated, in a classical way, through low-level unsupervised image segmentation of an input Magnetic Resonance (MR) image (from the Simple Linear Iterative Clustering proposed in [**3**] or the Graph-Based Image Segmentation technique described in [**56**]). Nevertheless, unlike MR images, which have a better quality, X-Ray images contain quantum noise, numerous artifacts (overlap and blurring) that can degrade any over-segmentation or superpixel map estimation procedure. In addition, the input image is over-segmented into small homogeneous regions or superpixels (sharing similar features) but which are not necessarily of interest in terms of bone or anatomical structures. Unlike this, our application generates the superpixel map by taking the intersection operation on the warped training label images which were manually segmented into different ROIs by experts. That is why our superpixel map takes into account the local and nonlinear deformations of ROIs (i.e., bone regions), existing in the training atlases, without too much error. Besides, our superpixel map is not used as an additional matching constraint for establishing the voxel-wise label transfer and fusion (as proposed in the two above mentioned references) but in order to refine the segmentation process.

As mentioned in our preliminary work [**122**] and to our knowledge, there is no reported work that proposes and uses, for the *label propagation* step, the result of a median partition-based optimization problem or that exploits a superpixel strategy for taking into account the nonlinear variability of the ROIs in label multi-atlas population. The remainder of this paper is divided into the following sections. In Section 2, we introduce details about the dataset that will be used in our model. Section 3 describes the DCT-based pre-processing step made on each X-ray image to be segmented in order to enhance the contour of the different anatomical structures to be segmented. The proposed model with its five steps, namely (1)

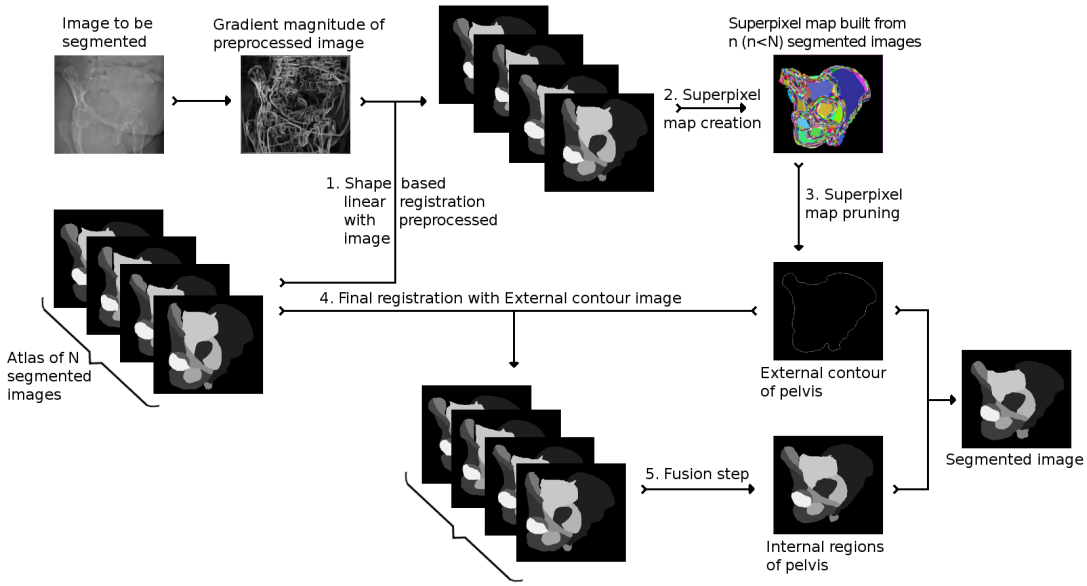


Fig. 3. The pipeline of the proposed approach.

shape-based linear registration and multi-atlas selection, (2) superpixel map creation, (3) superpixel map pruning, (4) final selection and filtering of the multi-atlas dataset, and (5) entropy-based fusion procedure are presented in Section 4. Finally, we show experimental results in Section 6.2 and conclude in Section 7. The pipeline of superpixel and multi-atlas segmentation of bone structures for the pelvic structure is given in Fig. 3.

2. Dataset And Multi-Atlas Creation

For our experiments (approved by the local ethics board of our University), we used 31 anonymized 45-degree oblique X-ray images of the pelvis, talus, and patella that were manually segmented into different ROIs by experts well trained in lower extremity and medical image segmentation. Images of the pelvis were partitioned into 14 different ROIs including the whole pelvis with sacrum, and both hemipelvis. Images of the talus were manually segmented into 4 different ROIs, including the body, dome and head of talus. of the patella had ROIs. (see Fig. 4)

3. X-ray Image Pre-processing

In our application, a pre-processing step is required in order to enhance the visibility of the bone contours which, in fact, constitute the most important and reliable low-level visual cue in each radiographic image. To this end, a histogram equalization technique and a DCT-based denoising step [**113, 174**] was used.

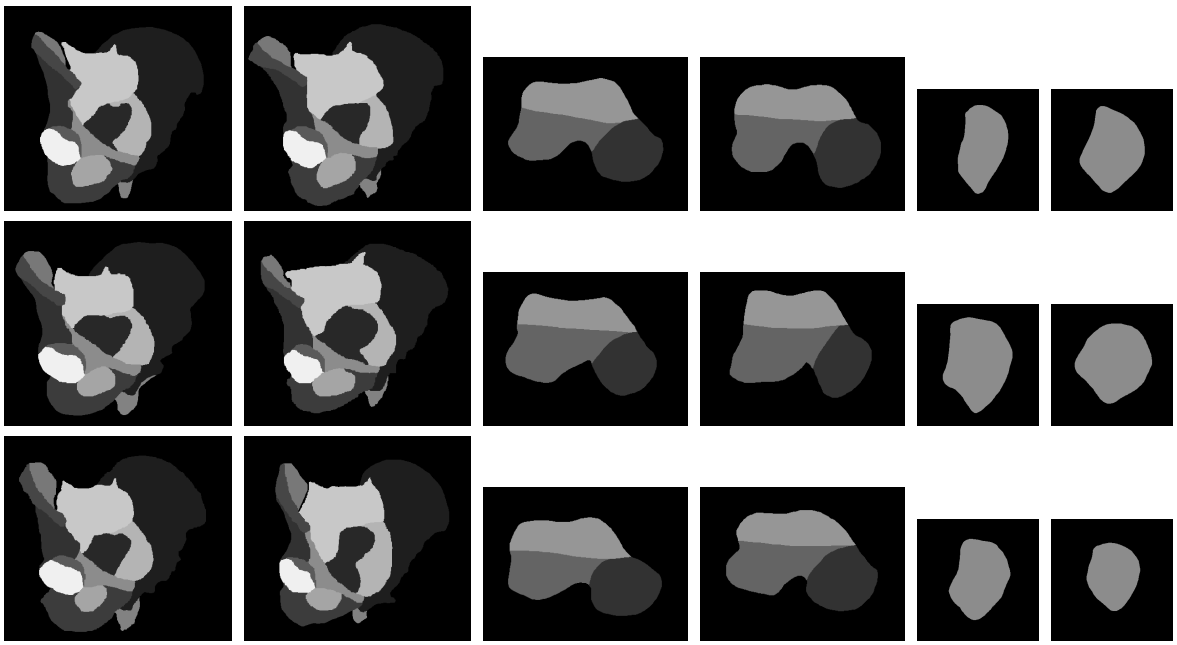


Fig. 4. Six examples of 45-degree oblique manual segmentations of the human pelvis (a), talus (b), and patella (c) used in our multi-atlas.

The DCT-based denoising procedure used in this application is an iterative strategy of denoising. Each iteration progressively denoises the image and consists in repeating a translation invariant DCT-based frequential filtering using a simple adaptive hard thresholding rule within a sliding subimage. Algorithmically and, more precisely, the iterative denoising scheme thus simply consists of alternating, until a maximal number of iterations is reached, two steps (which are also summarized in pseudo-code in Algorithm 1):

(1) The first step is a frequential filtering using the DCT transform of each 8×8 subimage extracted from the current denoised version of the degraded image (initially, this current image estimate is y, the degraded image itself). It is well known by the compression community that a DCT filtering (*i.e.*, a threshold on these DCT coefficients) on each 8×8 blocks extracted from a (noisy or not) image can denoise, but also can cause blocky artifacts on the processed image after reconstruction (*i.e.*, inverse DCT). In order to reduce these artifacts, a standard way is to use the DCT of all circularly translated version of the image, herein assumed to be toroidal. This thus implies to compute a set of $N_{\mathrm{S}} = 8 \times 8 = 64$ (8 horizontal shifts and 8 vertical shifts) transformed images (which are then averaged in the second step). For the filtering operation in the DCT domain, we used the simple hard thresholding rule classically used in the wavelet-based denoising approaches. After an inverse DCT transform, we have to un-shift the filtered image and store the result for the second step.

Algorithm 1 DCT-Based Denoising Algorithm

y Input image to be denoised (Input)
$\hat{x}$ Denoised estimated image (Output)
T Threshold

1. Initialization
$\hat{x}^{[1]} \leftarrow y$

2.
for iteration $n = 1$ to 5 **do**
I. Frequential Filtering
for All (8 horizontal and 8 vertical) possible shifts of $\hat{x}^{[n]}$ **do**
for All [8×8] blocks extracted from $\hat{x}^{[n]}$ **do**
• DCT Transform
• Threshold the obtained DCT coefficients w with the hard thresholding rule

$$\lambda_T^{\text{hard}} = \begin{cases} 0 & \text{if} \quad |w| \leq T \\ w & \text{otherwise} \end{cases}$$

• Inverse DCT Transform of these thresholded coefficients
end for
Unshift the filtered image and store it
end for
II. Averaging
$\hat{x}^{[n+1]} \leftarrow$ Averaging of these 64 denoised images
end for
$\hat{x} \leftarrow \hat{x}^{[n+1]}$

(2) The second operation consists in doing an averaging of these 64 denoised image estimates in order to compute the current denoised estimate and to make translation-invariant the proposed denoising procedure.

We returned to the first step for another denoising operation by considering that $y = \hat{x}$ and so on, until the stability of the solution, typically at the end of 5 full iterations of our application. For the implementation, we have used the fast 8×8, 16×16 FFT2D DCT package implemented in C code by T. Ooura (functions DDCT8X8S and DDCT16X16S tested in program SHRTDCT.C) and available on-line at http address given in [**125**].

4. Proposed Multi-Atlas Segmentation Model

The proposed multi-atlas segmentation process is mainly divided into two stages. The first part allows us to estimate the external contour of the bone structure to be segmented (steps 4.1, 4.2 and 4.3) and the second part (steps 4.4 and 4.5) is dedicated to the refinement of the external contour of the bone structure and, above all, to estimate the region labels within the bone structure to be semantically segmented.

Algorithm 2 Iterative Pruning Algorithm

S_{SupPix}	Superpixel map (Input)
I_{Grad}	Gradient potential field of the pre-processed X-ray image to be segmented (Input)
$\hat{S}_{\text{ExtCont}}$	External contour of the bone (Output)
$N_{\max}$	Maximal number of iterations
c_i	Thresholds ($i \in \{1,2,3\}$)
E_{ExtCont}	Energy of the external contour

1. Initialization

$\hat{S}^{[0]}_{\text{SPixPruned}} \leftarrow S_{\text{SupPix}}$ & $n \leftarrow 0$ & $E^{[0]}_{\text{ExtCont}} \leftarrow 0$

• Estimation of the biggest (usually central) superpixel *bsp* (which should never be pruned)

2. Iterative Pruning

repeat

for each superpixel *sp* ($sp \neq bsp$) of $\hat{S}^{[n]}_{\text{SPixPruned}}$, both in the lexicographic order and belonging to the external contour **do**

• Compute energy E_{Sp} of *sp* (using I_{Grad})

if ($E_{\text{Sp}} < c_2$) **then**

Delete *sp* from $\hat{S}^{[n]}_{\text{SPixPruned}}$

else

• Compute E_1, E_2: the energy of the external contour of $\hat{S}^{[n]}_{\text{SPixPruned}}$ with & without *sp* (using I_{Grad})

• Compute E_3, E_4: the energy of the external contour of *sp* and the energy of the superpixel which replaces *sp* (in the external contour of $\hat{S}^{[n]}_{\text{SPixPruned}}$) (using I_{Grad})

if ($E_2 \geq c_3 \cdot E_1$) and ($E_4 \geq c_3 \cdot E_3$) **then**

Delete *sp* from $\hat{S}^{[n]}_{\text{SPixPruned}}$

end if

end if

end for

• Compute energy $E^{[n]}_{\text{ExtCont}}$ (using I_{Grad})

$n \leftarrow n + 1$

until $\left(\hat{S}^{[n]}_{\text{SPixPruned}} \neq \hat{S}^{[n-1]}_{\text{SPixPruned}}\right)$ and $\left(n < N_{\max}\right)$ and $\left(E^{[n]}_{\text{ExtCont}} \geq c_1 \cdot E^{[n-1]}_{\text{ExtCont}}\right)$

$\hat{S}_{\text{ExtCont}} \leftarrow$ External contour of $\hat{S}^{[n]}_{\text{SPixPruned}}$

4.1. Shape-Based Linear Registration and Multi-Atlas Selection Step

Since the most significant and informative visual information of a bone structure in an X-ray image (especially in the lower limbs) remains its external bone contour, each segmented image of the multi-atlas was first independently and rigid-registered, through affine transformations, to the target image using a contour-based registration technique. This rigid registration was done to produce an initial alignment for each multi-atlas segmented image that maximized a similarity measure between an edge potential field estimated from the

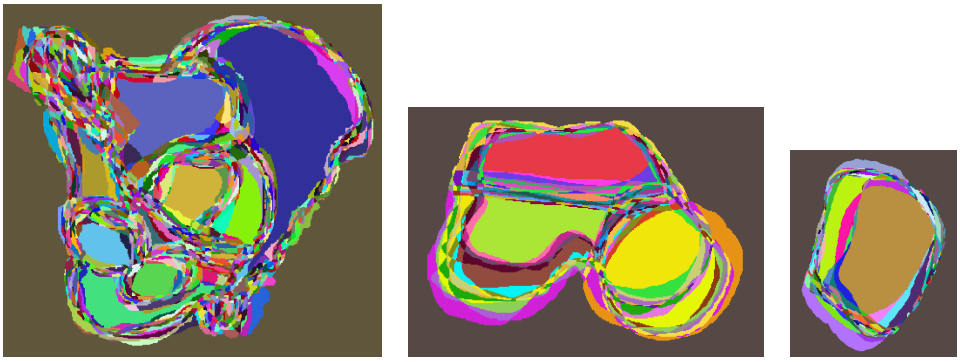


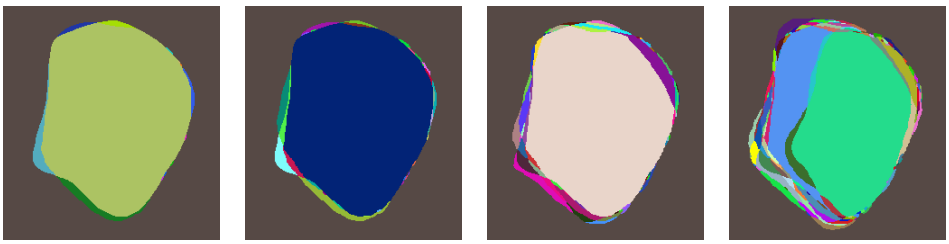
Fig. 5. From left to right; superpixel map obtained for a given X-ray 45 degree image for respectively the pelvis, talus, and patella to be segmented.

pre-processed target X-ray image and the external contour of the corresponding bone region in the segmented image. The edge potential was simply generated by the calculation of the gradient magnitude convolved with a Gaussian kernel (Gaussian blur) with a variance controlling the degree of smoothness of this contour field [**20**] and making nearly convex the energy function used in the optimization-based registration technique (see Fig. 9). The rigid registration that maximized the summation of this potential map over all the pixels located on the boundary of the external contour of the bone structure to be segmented (given by the manual segmentation) thus allowed us to get a rough alignment of each atlas segmented image to the target image.

Furthermore, our contour-based similarity registration score was also used to rank the set of rigid-registered manual segmentations of our multi-atlas by decreasing order of similarity. The first half of this rigid-registered segmentations was then used for estimating the adaptive superpixel representation associated to each input image to be segmented.

4.2. Superpixel Map Creation

A set of superpixels is thus estimated from the first half of the optimally rigidly registered segmented images belonging to the multi-atlas by simply taking the intersection of all the segments (or regions) existing in this pre-selected (rigid-registered) manual segmentations (see Fig. 5). Conceptually, each generated superpixel corresponds to a group of connected pixels belonging to the same region label and the set of superpixels provides a (superpixel) map which is assumed to divide the input image to be segmented into atomic regions sharing the same label (thus reducing the number of entities to be labeled or/and to guide and adding constraints to the segmentation process). Fig. 6 illustrates different superpixel maps of patella with different number of input atlases. Each superpixel (i.e., each color) represents a local and nonlinear deformations of ROIs of the patella.



(a) (b) (c) (d)

Fig. 6. Examples of superpixel maps of the patella created from (a) 2 atlases, (b) 3 atlases, (c) 5 atlases and (d) 10 atlases. Each superpixel (i.e. each color) represents a nonlinear deformation between each bone regions.

Ren and Malik first used the superpixels [**130**], as a pre-processing step, for the segmentation of natural images in order to simplify and better spatially constrain the segmentation task. This concept has also been proposed in [**77**] however, with the different purpose of combining a segmentation map (into spatial regions) and a coarsely estimated and to-be-refined higher level vision application (e.g., occlusion map, motion segmentation/estimation field, etc.). In multi-atlas segmentation, the concept of superpixels have also been used, but especially for MR images in [**167, 175**] and in order to oversegment the target images and all atlases into small regions, sharing similar features but which are not necessarily of interest in terms of bone structures. In our application, the superpixel maps were generated directly from the well segmented training atlases and were thus adaptive for each radiographic image to be segmented. They will allow us both to carry out an adaptive local non-rigid registration taking into account the local and nonlinear deformations of ROIs specific to the target but also to guide and to simplify the superpixel map pruning step by better spatially constraining it.

4.3. Superpixel Map Pruning

At this stage, an adaptive superpixel map, which non-parametrically incorporated all the nonlinear variability existing in a selected subset of the multi-atlas dataset (inter-patient variability of the imaging pose and bone morphology), was generated using each radiographic target image to be segmented. Since the superpixel map was already linearly registered to the target image (because this intersection map was built from a subset of selected and linearly rigidly registered segmentations), a pruning step, using an iterative pruning algorithm, was easily achieved by simply finding the set of connected superpixels which maximized the contour-based similarity between the outer contour of this superpixel map and the edge map of the radiographic image[1]. More precisely, this iterative pruning procedure can be

[1]Experiments have shown that slightly better results are given if the superpixel map is scaled by a factor slightly greater than one in order to ensure that the external bone contour of the X-ray target image is fully

described in the following manner; Each superpixel connected with the background label (*i.e.*, which is contiguous with the outer bone region to be segmented), in a lexicographic order, was classified as belonging to the considered bone region if the outer contour-based similarity metric increased and until convergence or a maximum number of iterations was reached. At convergence, this iterative procedure (See Algorithm 2) allowed us to estimate a precise closed (and continuous) outer contour of the bone structure.

4.4. Final Selection/Filtering of the Multi-atlas Dataset

At this stage, we had to refine the previously closed external contour of the bone structure and also to infer both the internal regions (within the bone structure) along with their most likely semantic labels.

To this end, since the previously estimated external contour was closed, a final linear registration, between it and each manual segmentation of the multi-atlas, was achieved with a region-based similarity metric (namely the F-measure[2]) and also allowed us to identify and select the first half of the manual segmentations of our multi-atlas dataset, which were best registered.

Finally, these above-mentioned pre-selected rigid-registered manual segmentations were combined in the variation of information (VoI) sense [**109**], to yield the final semantic segmentation.

4.5. VoI-based Fusion Procedure

The fusion procedure, used in our application, is based on the variation of information (VoI) distance [**103, 104**] which has the crucial advantage of being a true metric on the space of segmentations, *i.e.*, a metric which is positive, symmetric and obeying the triangle inequality [**104**]. This metric is commonly used for comparing the similarity of two segmentation results (or clusterings) or also as an efficient quality score that measures the agreement of the segmentation result with a given ground truth. To this end, it was used in image segmentation [**105, 106, 110**] in order to objectively benchmark the efficiency of an unsupervised segmentation algorithm giving an automatic machine segmentation of an image to a ground truth segmentation (e.g., a manually hand-segmented image given by an expert).

---

contained in the superpixel map. To this end, after trial and error, a scale factor of 1.02 allows us to give the best segmentation results.

[2]In this binary region-based classification step, the F-measure (also called F-score) is a region-based measure of accuracy. The accuracy of this classification is expressed by means of recall (R), precision (P) and their harmonic mean, the F-score: $F = 2\,P\,R/(P+R)$ with $P = TP/(TP+FP)$ and $R = TP/(TP+FN)$, where TP, FP and FN indicate, respectively, the number of correct detections, false alarms and missed detections at pixel-level. The F-score can be also interpreted as a weighted average of the precision and recall, where an F-score reaches its best value at 1 and worst at 0.

Let $S^{\mathrm{A}} = \{C_1^{\mathrm{A}},C_2^{\mathrm{A}},\ldots,C_{R^{\mathrm{A}}}^{\mathrm{A}}\}$ and $S^{\mathrm{B}} = \{C_1^{\mathrm{B}},C_2^{\mathrm{B}},\ldots,C_{R^{\mathrm{B}}}^{\mathrm{B}}\}$ be respectively the first and second segmentation (or the machine segmentation result to be evaluated and the ground truth segmentation of the same size) between which the VoI distance has to be estimated. S^{A} is composed of n pixels and R^{A} regions C_i (also called cluster, segments or superpixels in the following). Let n_i^{A} be the number of pixels in the i-th region (C_i) of the segmentation S^{A}, n_j^{B} the number of pixels in the j-th region (C_j) of the segmentation S^{B} and finally n_{ij} the number of pixels which are together in the i-th region of the segmentation S^{A} and in the j-th region of the segmentation S^{B}. $P(i) = n_i^A/n$ represents the probability that a pixel belongs to the i-th region of S^{A} (respectively $P(j) = n_j/n$ being the probability that a pixel belongs to the j-th region of S^{B}). Let finally $P(i,j) = n_{ij}/n$ represents the joint probability that a pixel belongs to C_i^{A} and to C_j^{B}. The VoI between S^{A} and S^{B} is always in the range [0,...1] and defined as:

$$\begin{aligned}
\mathrm{VoI}(S^{\mathrm{A}},S^{\mathrm{B}}) &= H(S^{\mathrm{A}}) + H(S^{\mathrm{B}}) - 2\cdot I(S^{\mathrm{A}},S^{\mathrm{B}}) \\
\text{with:}\quad H(S^{\mathrm{A}}) &= -\sum_{i=1}^{R^{\mathrm{A}}} P(i)\,\log P(i) = -\sum_{i=1}^{R^{\mathrm{A}}} \frac{n_i^{\mathrm{A}}}{n}\,\log\frac{n_i^{\mathrm{A}}}{n} \\
H(S^{\mathrm{B}}) &= -\sum_{i=1}^{R^{\mathrm{B}}} P(j)\,\log P(j) = -\sum_{i=1}^{R^{\mathrm{B}}} \frac{n_j^{\mathrm{B}}}{n}\,\log\frac{n_j^{\mathrm{B}}}{n} \\
I(S^{\mathrm{A}},S^{\mathrm{B}}) &= \sum_{i}^{R^{\mathrm{A}}}\sum_{j}^{R^{\mathrm{B}}} P(i,j)\,\log\frac{P(i,j)}{P(i)\,P(j)}
\end{aligned} \tag{4.1}$$

where $H(S^{\mathrm{A}})$ and $H(S^{\mathrm{B}})$ denote respectively the classical entropy associated with the segmentation S^{A} and S^{B} and $I(S^{\mathrm{A}},S^{\mathrm{B}})$ the mutual information between these two segmentations.

Let $\{S_k\}_{k\leq L}$ be a finite ensemble of L segmentations $\{S_k\}_{k\leq L} = \{S_1,S_2,\ldots,S_L\}$ to be fused or combined in the VoI metric sense in order to obtain $\hat{S}_{\overline{\mathrm{VoI}}}$, *i.e.*, (equivalently) the average (mean) segmentation in the VoI distance sense also called the best compromise segmentation solution (resulting in a consensus in terms of contour accuracy, detail level, clutters, etc. exhibited by each segmentations ($\in \{S_k\}_{k\leq L}$)). $\hat{S}_{\overline{\mathrm{VoI}}}$ can be regarded as the average pairwise agreement between putative individual segmentations ($\in \{S_k\}_{k\leq L}$) or the solution of the following optimization (or so-called *median partition* [**163**]) problem:

$$\hat{S}_{\overline{\mathrm{VoI}}} \;=\; \arg\min_{S\in\mathcal{S}_n} \underbrace{\frac{1}{L}\sum_{k=1}^{L} \mathrm{VoI}\,(S,S_k)}_{\overline{\mathrm{VoI}}\,(S,\{S_k\}_{k\leq L})} \tag{4.2}$$

with $\mathcal{S}_n$ is the set of all possible segmentations using n pixels. This consensus segmentation solution can be expressed as the result of a minimization problem on the consensus function $\overline{\mathrm{VoI}}$ and can be efficiently solved with a iterative local gradient descent algorithm [**109**] in which a new label x is assigned to pixel s (initially with label l_s), if this pixel is connected

to the x-th region and if there is a local decrease in the energy $\overline{\mathrm{VoI}}(.)_{s:l_s \to x}$ with:

$$\begin{aligned}
&\Delta\overline{\mathrm{VoI}}\left(\hat{S}^{[p]}_{\overline{\mathrm{VoI}}}, \{S_k\}_{k\leq L}\right)_{s:m\to x} = \\
&\quad L \cdot \left\{ -\tfrac{n_m}{n}\log\left(\tfrac{n_m}{n}\right) - \tfrac{n_x}{n}\log\left(\tfrac{n_x}{n}\right) \right. \\
&\quad \left. + \tfrac{(n_m-1)}{n}\log\left(\tfrac{n_m-1}{n}\right) + \tfrac{(n_x+1)}{n}\log\left(\tfrac{n_x+1}{n}\right) \right\} \\
&\quad -2 \cdot \textstyle\sum_{l=1}^{L} \left\{ \tfrac{n_{m,\mathcal{L}_s^l}}{n}\log\left(\tfrac{n_{m\mathcal{L}_s^l}}{n} \cdot \tfrac{n}{n_m} \cdot \tfrac{n}{n_{\mathcal{L}_s^l}}\right) \right. \\
&\quad + \tfrac{n_{x\mathcal{L}_s^l}}{n}\log\left(\tfrac{n_{x\mathcal{L}_s^l}}{n} \cdot \tfrac{n}{n_x} \cdot \tfrac{n}{n_{\mathcal{L}_s^l}}\right) \\
&\quad - \tfrac{(n_{m\mathcal{L}_s^l}-1)}{n}\log\left(\tfrac{(n_{m\mathcal{L}_s^l}-1)}{n} \cdot \tfrac{n}{(n_m-1)} \cdot \tfrac{n}{n_{\mathcal{L}_s^l}}\right) \\
&\quad \left. - \tfrac{(n_{x\mathcal{L}_s^l}+1)}{n}\log\left(\tfrac{(n_{x\mathcal{L}_s^l}+1)}{n} \cdot \tfrac{n}{(n_x+1)} \cdot \tfrac{n}{n_{\mathcal{L}_s^l}}\right) \right\}
\end{aligned} \tag{4.3}$$

where $\mathcal{L}_s^l$ denotes the label at site s of the l-th segmentations ($l \leq L$) of the segmentation ensemble $\{S_k\}_{k\leq L}$ and we recall that $n_{m\mathcal{L}_s^l}$ designates the number of pixels which are together in the m-th region of the segmentation S and in the $\mathcal{L}_s^l$-th region of the segmentation $S_l \in \{S_k\}_{k\leq L}$ (see Algorithm 1). As initialization of this steepest gradient descent, we can start from the segmentation result (among the L segmentation results to be averaged), ensuring the minimal consensus energy in the $\overline{\mathrm{VoI}}$ sense [**109**].

In the *label propagation* step, the VoI-based fusion procedure allowed us both to refine the external contour of the bone structure to be segmented, but also, and especially, to infer the internal region labels from the selected atlases (previously estimated in Section 4.4) to the final segmentation map. Fig. 7 shows two examples of the evolution of the consensus energy function of the VoI-based fusion procedure along the number of iterations for two different bone structures.

5. Time Complexity

The purpose of this section is to compute an estimate for the time complexity required to segment an X-ray image by our strategy. We performed a complexity analysis on each step of our approach.

First, the time complexity of our X-ray image pre-processing step is linear with respect to the number of pixels with $\mathcal{O}(p(n_{dct}b + g))$ where p is the number of pixels in the image, n_{dct} is the number of DCT iterations, b is the extracted block size and g is the Gaussian kernel size.

Algorithm 3 VoI-Based Segmentation Fusion Algorithm

$\{S_k\}_{k\leq L}$	Set of L segmentations to be fused (Input)
$\hat{S}_{\overline{\text{VoI}}}$	Estimated fused result (Output)
$T_{\max}$	Maximal number of iterations
$\overline{\text{VoI}}$	Mean VoI (See Eq. (2))

1. Initialization

$$\hat{S}^{[0]}_{\overline{\text{VoI}}} = \underset{S\in\{S_k\}_{k\leq L}}{\operatorname{argmin}} \overline{\text{VoI}}\,(S,\{S_k\}_{k\leq L}) \quad \& \quad n \leftarrow 0$$

2. Steepest Local Energy Descent
while $n < T_{\max}$ **do**
  for each pixel of $\hat{S}^{[n]}_{\overline{\text{VoI}}}$ with label l_s at site s **do**
    • Let $\mathcal{E}$ the set of labels $\neq l_s$ contained in the local (squared) fixed-size ($N_w = 7$) neighborhood of s
    • Draw a new label x according to the uniform distribution in the set $\mathcal{E}$
    if $x = \emptyset$ **then**
      continue;
    end if
    if pixel s with label x is not 4-connected with the x-th region in $\hat{S}^{[n]}_{\overline{\text{VoI}}}$ **then**
      continue;
    end if
    • Compute $\Delta\overline{\text{VoI}}\left(\hat{S}^{[n]}_{\overline{\text{VoI}}}, \{S_k\}_{k\leq L}\right)_{s:l_s\to x}$ (See Equation (3))
    if $\Delta\overline{\text{VoI}}\left(\hat{S}^{[n]}_{\overline{\text{VoI}}}, \{S_k\}_{k\leq L}\right)_{s:l_s\to x} > 0$ **then**
      replace label l_s by label x at site s
    end if
  end for
  $n \leftarrow n+1$
end while
$\hat{S}_{\overline{\text{VoI}}} \leftarrow \hat{S}^{[n+1]}_{\overline{\text{VoI}}}$

Second, the time complexity of bone structure external contour estimation step is $\mathcal{O}(p(N_{dataset}ts + n_{pruning}P))$. More precisely, the time complexity of shape-based registration is $\mathcal{O}(N_{dataset}tsp)$, where $N_{dataset}$ is the number of images in training dataset, t is the number of translations, and s is the number of scaling factors. This complexity grows linearly with $N_{dataset}$ and p, while t and s can be fixed. The time complexity of superpixel map creation is $\mathcal{O}(N_{map}p)$, where N_{map} ($N_{map} < N_{dataset}$) is the number of images used for creating the superpixel map. Also, the time complexity of the superpixel map pruning algorithm is $\mathcal{O}(n_{pruning}Pp)$, where $n_{pruning}$ is the number of iterations required and P is the number of superpixels.

Third, the time complexity of external contour refinement and internal segmentation estimation step is $\mathcal{O}(p(N_{dataset}ts + n_{voi}N_{fusion}))$. More precisely, the time complexity of

Algorithm 4 The proposed X-ray segmentation algorithm.

I	X-ray image to be segmented (Input)
$\{\mathcal{S}_k\}_{k\leq N}$	DataSet (multi-atlas) of N (*a priori*) segmentations of the bone structure (Input)
I_{Seg}	Segmented X-ray image (Output)
T_{DCT}	Threshold of the denoising filter
σ_b	Std of the contour-based potential field blur
ρ	Selection % of the best registrations
$\{c_i\}_{i=1,2,3}$	Thresholds of the pruning algorithm

Initialization: X-Ray Image Pre-processing

• **From** I **do**

Histogram equalization

– DCT-based denoising algorithm (T_{DCT})

– Gradient potential field (gradient magnitude + Gaussian blur with standard deviation σ_b)

• **Result:** I_{Grad}

1. Bone Structure External Contour Estimation

• A. Shape-based Linear Registration and Multi-Atlas Selection

$\{\mathcal{S}_k^\bullet\} \leftarrow$ Rigid registration of each segmentation of $\{\mathcal{S}_k\}_{k\leq N}$ and selection of the ρ% best registrations (in term of our contour-based potential using I_{Grad})

• B. Superpixel Map Creation

$S_{\text{SupPix}} \leftarrow$ intersection of all the regions existing in $\{\mathcal{S}_k^\bullet\}$

• C. Superpixel Map Pruning

$\hat{S}_{\text{ExtCont}} \leftarrow$ Pruning Algorithm $\left(S_{\text{SupPix}}, I_{\text{Grad}}, \{c_i\}_{i=1,2,3}\right)$

2. External Contour Refinement and Internal Segmentation Estimation

• D. Final Selection of the Multi-Atlas

$\{\mathcal{S}_k^\circ\} \leftarrow$ Rigid registration of each segmentation of $\{\mathcal{S}_k\}_{k\leq N}$ in the (region-based) F-score sense between $\hat{S}_{\text{ExtCont}}$ and $\mathcal{S}_k$ & selection of the ρ% best registrations

• E. VoI-Based Mean Segmentation Estimation

$I_{\text{Seg}} \leftarrow$ Voi-Based Mean Seg. Algorithm $\left(\{\mathcal{S}_k^\circ\}\right)$

final selection of the multi-atlas (*i.e.* rigid registration) step, which is similar to the shape-based registration, is $\mathcal{O}(N_{dataset}tsp)$. Also the time complexity of the VOI based fusion algorithm is $\mathcal{O}(n_{voi}N_{fusion}p)$, where n_{voi} is the number of iterations required, and N_{fusion} ($N_{fusion} < N_{dataset}$) is the number of images used for fusion step.

Finally, the overall complexity of our approach remains linear with respect to the number p of pixels in the image with $\mathcal{O}(p(n_{dct}b + g + N_{dataset}ts + n_{pruning}P + n_{voi}N_{fusion}))$.

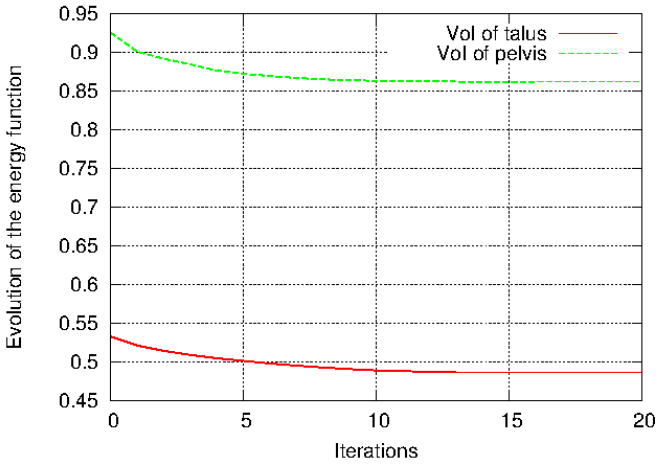


Fig. 7. Evolution of the consensus energy function of the VoI-based fusion procedure along the number of iterations for two different bone structures.

6. Experiments

6.1. Data

For the experiments, we validated our multi-atlas based segmentation approach on 31 45-degree oblique X-ray radiographic images and on three different human bone structures (pelvis, talus, and patella) acquired using a X-ray imaging system. Every pelvis, talus and patella image was cropped for the best view, resulting in image sizes of 1500×1350, 400×300 and 350×350 pixels respectively.

6.2. Experimental Results

The overall algorithm is summarized in pseudo-code in Algorithm 4. The internal parameters of the proposed method are respectively the threshold T_{DCT} of the DCT-denoising filter for the pre-processing step, the standard deviation σ_b of the potential field used in the shape-based linear registration, and local non rigid contour based reconstruction (step 1.A and 1.C). The selection parameter ρ of the multi-atlas selection (step 1.A and 2.A), and finally the three thresholds of the pruning algorithm c_1, c_2, c_3. We have set these parameters, chosen after some trial and error experiments, to the following respective values $T_{\mathrm{DCT}} = 40$, $\sigma_b = 2$, $\rho = 50\%$ and $c_1 = 0.95$,$c_2 = 3.5$,$c_3 = 0.99$.

The performance of segmentation algorithms is usually evaluated based both on the quality or accuracy of the segmentation result and computational requirement. The quality or accuracy of the segmentation algorithm is evaluated by comparing the output of the method with a ground truth, usually obtained from manual segmentation done by experienced radiologists. However, the ground truth always suffers from inter and intra-observer variations.

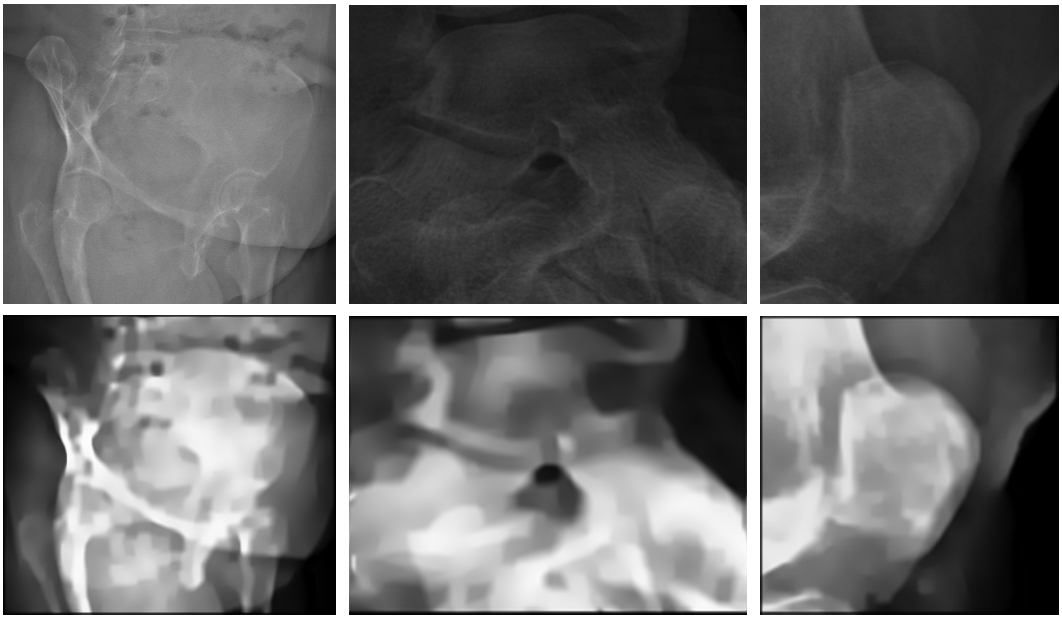


Fig. 8. Original oblique X-ray radiographic images of respectively the pelvis, the talus and the patella before and after the denoising step (see Section 3).

To solve this problem, we have performed a leave-one-out procedure, *i.e.,* we removed each existing manual segmentation from the multi-atlas data set while other manual segmentations remained. Each X-ray image associated to the previously removed segmentation map was then segmented by our strategy and compared, in terms of classification error rate (i.e., similarity index), with its manual segmentation.

Step	**Time (sec.) and %**	
Pre-processing	1	(0.5%)
Shape-based linear registration	4	(2.5%)
Superpixel map estimation	2	(1%)
Iterative pruning algorithm	15	(9%)
External contour-based registration	44	(27%)
VoI-based mean segmentation	95	(60%)
Total computational time	161	(100%)

Tableau 2. Computational time of different steps of the algorithm for the pelvis.

Fig. 8 shows the original oblique X-ray radiographic image of pelvis, talus and patella respectively before and after the pre-processing denoising step (without the computation of the magnitude gradient) (see Section 3). Fig. 9 shows an example of gradient magnitude and potential field with Gaussian filter with an increasing standard deviation σ_b of three original oblique pre-processing X-ray radiographic images of respectively the pelvis, talus, and patella. Fig. 10 shows of the external pelvic, talar, and patellar contours respectively on the corresponding 45-degree oblique X-ray radiographic (original and gradient) images and Fig. 11 shows the resulting segmentations estimated after the VoI-based *label propagation*

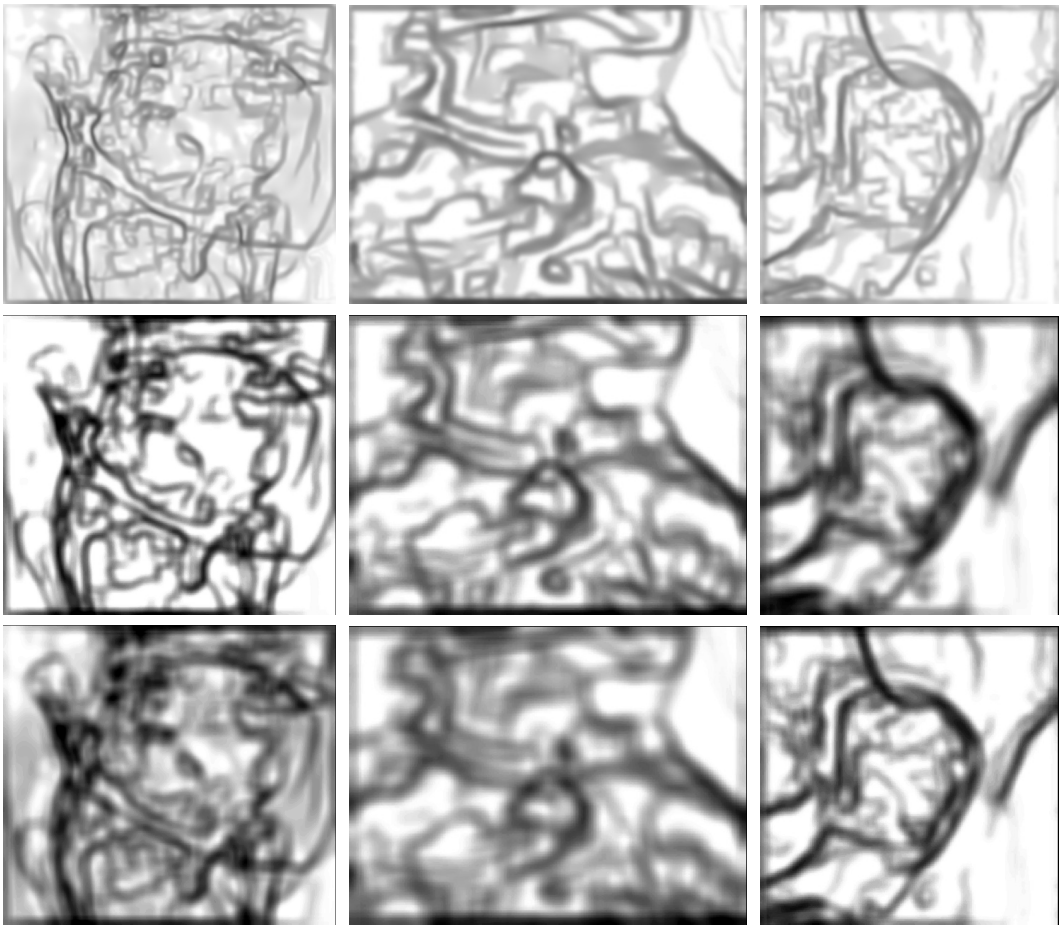


Fig. 9. From left to right; gradient magnitude of three original oblique pre-processing X-ray radiographic images of the pelvis, talus, and the patella, respectively. From top to bottom; potential field with Gaussian filter with an increasing standard deviation σ_b.

step, with the estimated internal region labels of the pelvic, talar and patellar structures. Figures 12, 13, and 14 show respectively some segmentation results from our automatic segmentation approach compared to a manual gold standard segmentation.

The accuracy of our automatic segmentation approach was, with our leave-one-out procedure, respectively 85.02% for the pelvis (e.g. in Fig. 12), 88.3% for the talus (e.g. in Fig. 13), and 93.79% for the patella (e.g. in Fig. 14). These scores take into account the accuracy of the labeling of the different regions within the bone structure to be segmented. The complete segmentation of one image took, on average, approximately 161 seconds for the pelvis, 62 seconders for the talus, and 80 seconds for the patella on a 64-bit desktop PC (2,50GHz Core i5-3210M CPU and a graphic card with Intel, 8GB RAM). Multi-core parallel computing with Open MP was used during the rigid registration step for each individual registration to speed up the overall procedure, Table 2 shows the running time of the different steps of our segmentation approach.

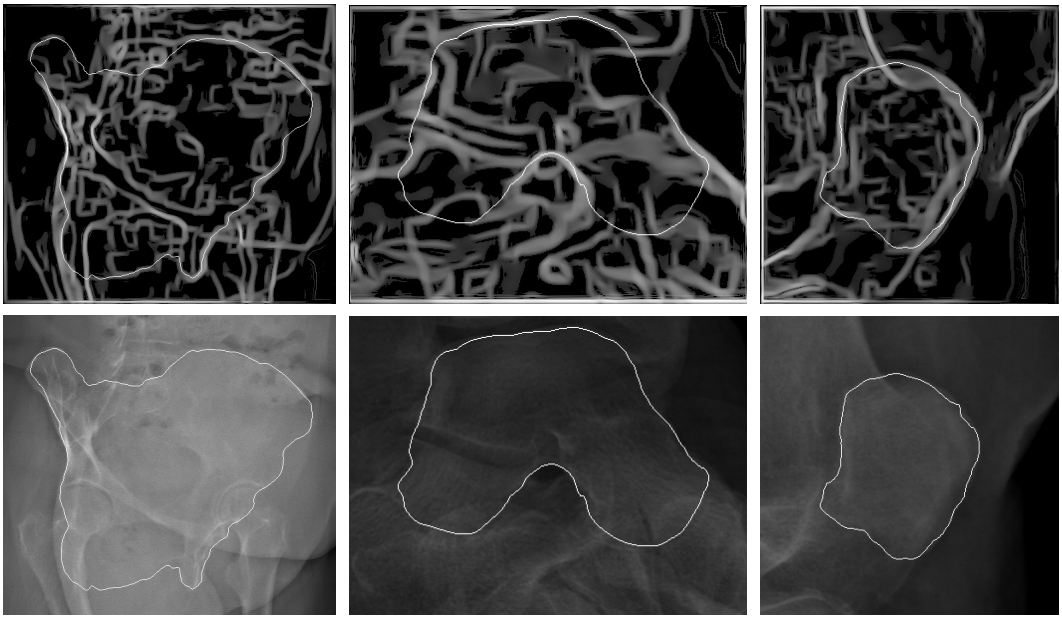


Fig. 10. Respectively the pelvic, talar and patellar external contours on the 45 degree oblique X-ray radiographic images (gradient and original).

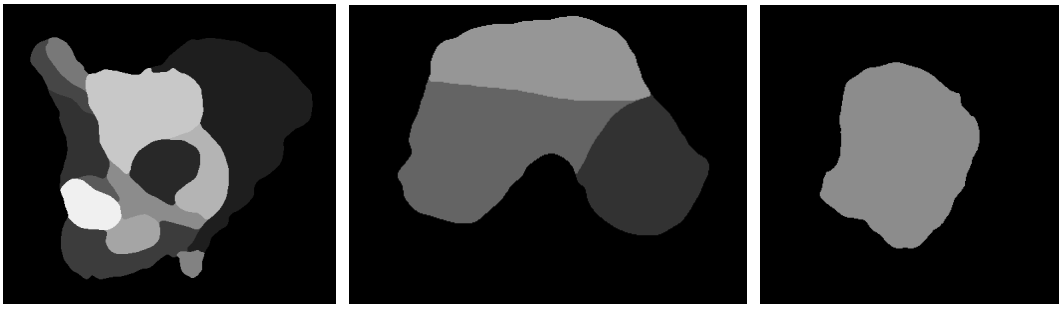


Fig. 11. Resulting fusion image estimated after the VoI-based *label propagation* step, with the estimated internal region labels of the pelvis, talus, and patella.

In spite of the numerous artifacts existing in X-ray images, such as the overlapping edges and concealed boundaries, the proposed segmentation method can cope with these artifacts and achieve good segmentation performance in an unsupervised way.

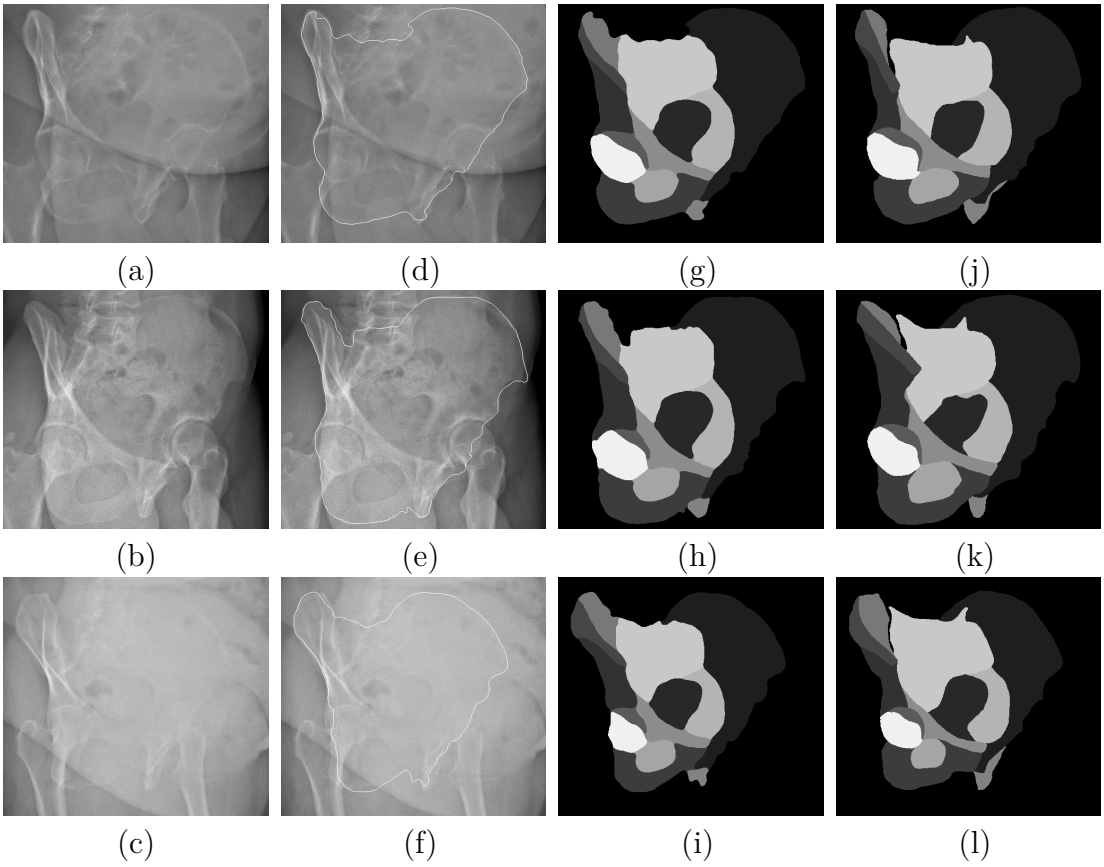


Fig. 12. Comparison of segmentation results from our approach and a manual segmentation of the pelvis. (a), (b), and (c) original oblique X-ray radiographic images; (d), (e), and (f) external pelvic contours on the corresponding X-ray images; (g), (h), and (i) resulting images; (j), (k), and (l) manual segmentations.

6.3. Comparison with other methods

We compared our method to the probabilistic patch-based label fusion model and the classical patch-based majority voting fusion scheme[3] using different registration strategies. Briefly, as in [**11**], atlases and corresponding label maps were first warped to the target image using image registration. Then, in order to determine the label at a given voxel, each atlas patch in the target image was compared to a number of patches (in a local search region) in a possibly warped atlases and was then weighted according to its similarity and distance to the target patch. Finally, labels from all the atlas patches were combined by a weighted majority voting based strategy, to give the label estimate in the target image. The resulting

[3]Patch-based majority voting is a powerful tool and also the simplest label fusion method exploited in multi-atlas based segmentation. It is based on the assumption that, if patches of the input subject image are locally similar to the patches of atlases, they should have a similar label. To this end, the more similar patches are selected and the label in the target image can then be determined by choosing the majority vote, *i.e.,* the label which is the most frequent at this voxel [**73**].

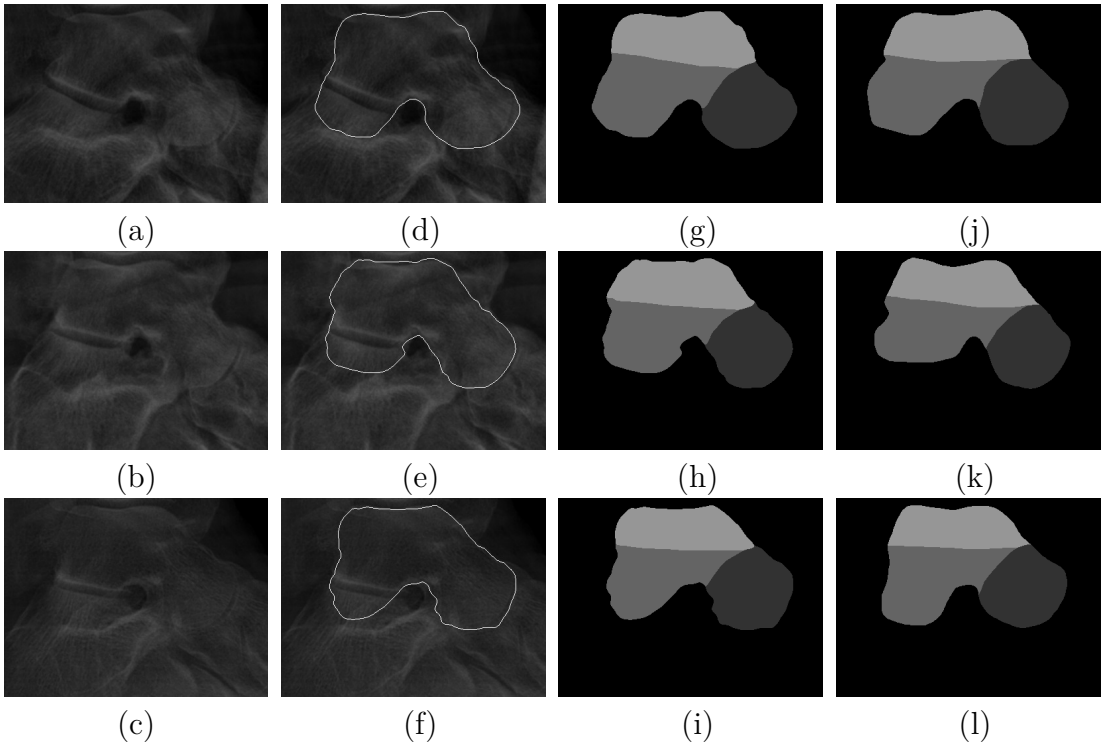


Fig. 13. Comparison of segmentation results from our approach and a manual segmentation of the talus. (a), (b), and (c) original oblique X-ray radiographic images; (d), (e), and (f) external talar contours on the corresponding X-ray images; (g), (h), and (i) resulting images; (j), (k), and (l) manual segmentations.

segmentation was then incorporated back into the image registration process to refine the registration results. We also compared our VoI based fusion step to Shape-based averaging (SBA)[4] fusion step presented in [**134**]. These two fusions of segmentations are based on a different similarity metric between two segmentation maps (i.e., the entropy concept for the VoI and a distance, related to the Chamfer distance for the SBA).

This comparison was evaluated on 31 radiographic image datasets for each bone structure and performed on above-mentioned desktop PC. In this comparison, we used an affine registration (AR) and our shape-based registration (SBR) to align the atlas with the target, then the label fusion process (MV, PB, SBA and our method VoI) was applied to obtain the final result. The segmentation quality was estimated with the similarity index by comparing the manual segmentations (considered as a gold standard) with the combinations of different registration and label fusion strategies (MV-AR, MV-SBR, PB-AR, PB-SBR, our method with VoI, and our method with SBA). For two binary segmentations A and B, the similarity index is computed as the ratio between $2|A \bigcap B|$ et $|A| + |B|$. Figures 15, 16, and 17 show respectively some segmentation results from these 6 abovementioned segmentation

[4]SBA is based on the distances from each pixel to the contours of each label, and thus, it implicitly includes the neighborhood information of pixels [**134**].

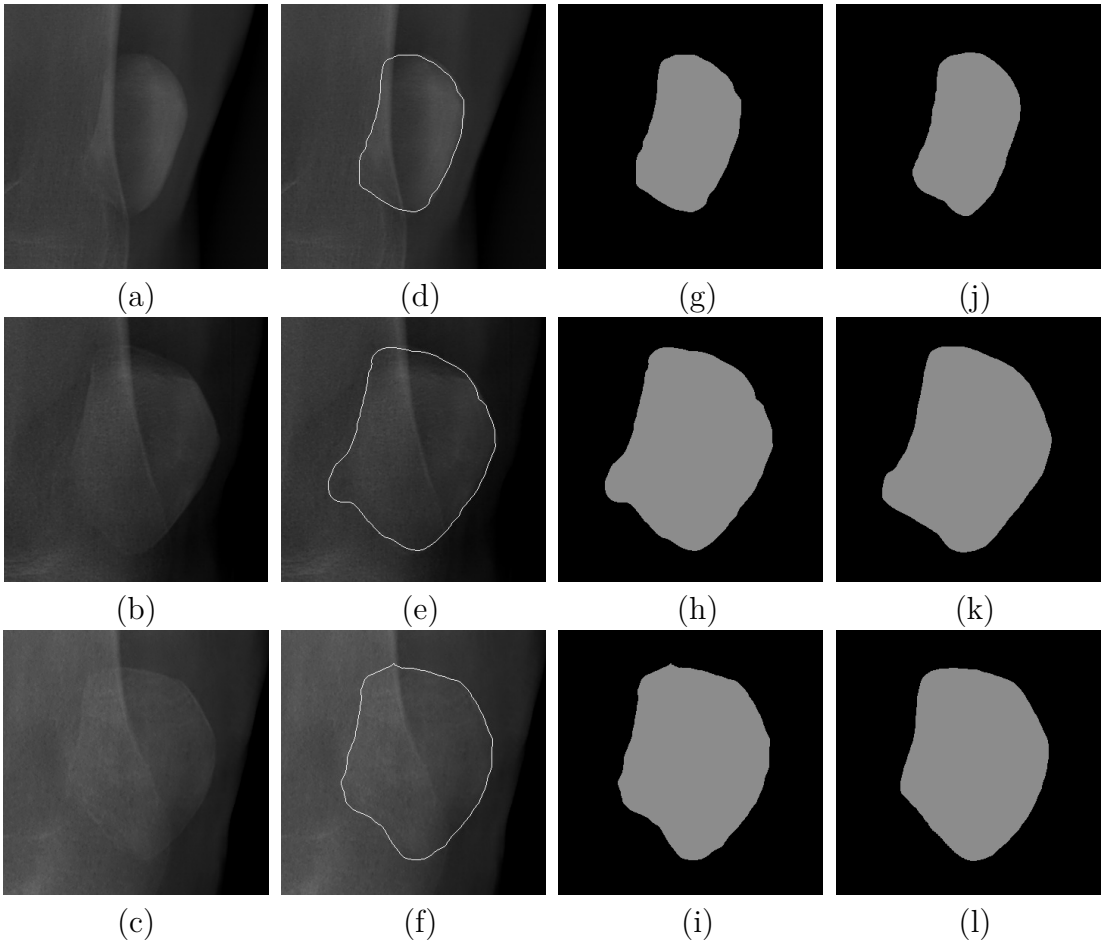


Fig. 14. Comparison of segmentation results from our approach and a manual segmentation of the patella. (a), (b), and (c) original oblique X-ray radiographic images; (d), (e), and (f) external patellar contours on the corresponding X-ray images; (g), (h), and (i) resulting images; (j), (k), and (l) manual segmentations.

approaches compared to a manual gold standard segmentation. The Table 3 and 4 show the comparison of the similarity index (*i.e.*, accuracy) and the computational time of these three segmentation methods respectively. Fig. 18 shows the accuracy of each atlas in dataset for leave-one-out validation.

		Patella	**Talus**	**Pelvis**
Our method	VoI based fusion	93.79%*	88.3%	85.02%
	SBA fusion		88.38%	85.01%
Majority voting	affine registration	89.99%	85.36%	67.77%
	shape-based registration	90.26%	85.26%	76.91%
Probabilistic patch-based	affine registration	90.01%	85.41%	68.03%
	shape-based registration	90.27%	85.44%	77.59%

Tableau 3. Accuracy of our method versus the MV and the PB fusion model with different registration and fusion methods for multi-atlas segmentation (*:Patella bone has only 1 ROI).

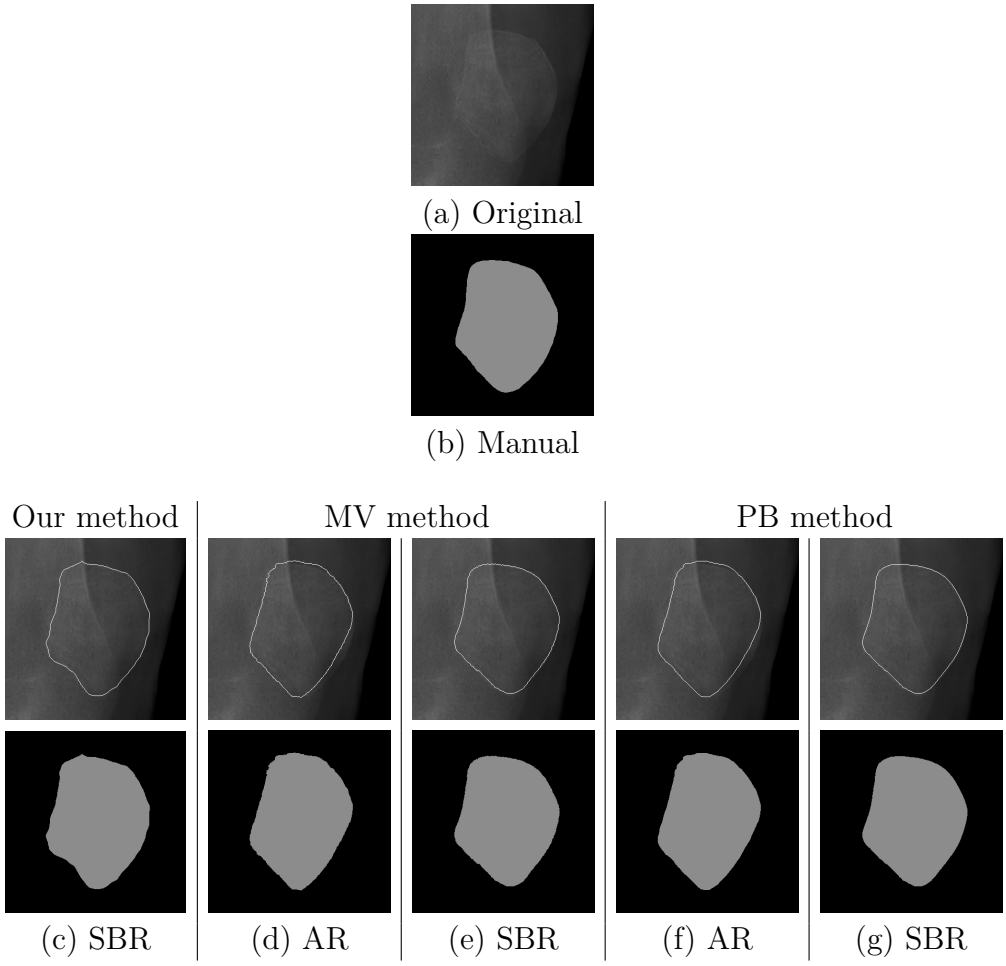


Fig. 15. Comparison of segmentation results of the patella. (a) original oblique X-ray radiographic image, (b) manual segmentation, third row: external patellar contours on the corresponding X-ray image, fourth row: resulting images.

		Patella	**Talus**	**Pelvis**
Our method	VoI based fusion	80*	62	161
	SBA fusion		52	129
Majority voting	affine registration	176	165	212
	shape-based registration	56	6	9
Probabilistic patch-based	affine registration	181	170	235
	shape-based registration	58	12	18

Tableau 4. Computational time (sec.) of our method versus the MV and the PB fusion model with different registration and fusion methods for multi-atlas segmentation (*:Patella bone has only 1 ROI).

7. Discussion and Conclusion

In our experiments, we observed that the VoI based fusion and SBA fusion have almost the same accuracy, probably due to the small number of ROIs. The experiments also showed that our shape-based registration gave a better result than image registration in the

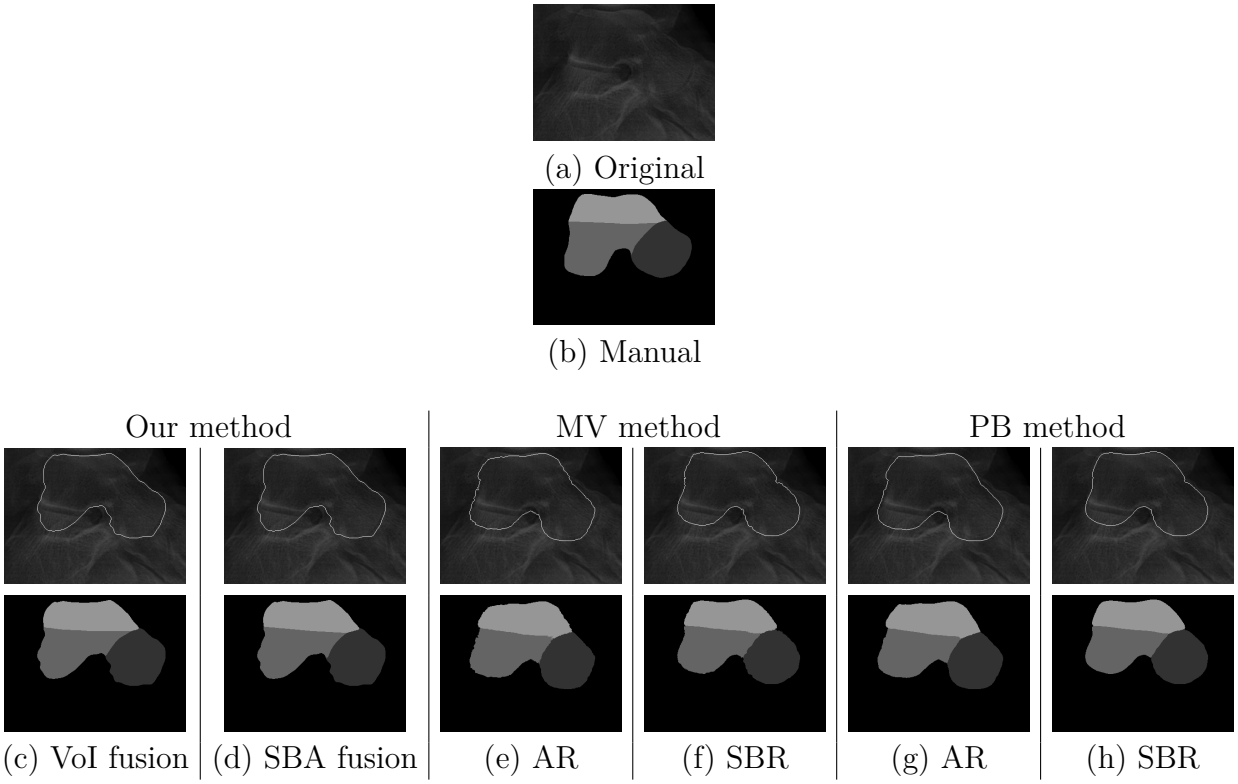


Fig. 16. Comparison of segmentation results of the talus. (a) original oblique X-ray radiographic image, (b) manual segmentation, third row: external talar contours on the corresponding X-ray image, fourth row: resulting images.

probabilistic patch-based label fusion model and the classical majority voting fusion scheme. Our method is slightly better than the other methods (MV and PB) with two different registration strategies in case of simple bones (patella, talus), and much better in the case of a complex structure (pelvis). Note that probabilistic patch-based label fusion methods are used for MRI, and our method is proposed for X-ray radiographic imaging. As already mentioned in the introduction, image registration and patch-based methods probably turn out not to be well suited for X-ray images since quantum noise can be different for different X-ray imaging systems and also because numerous artifacts exist in these images, mainly due to the fact that neighboring tissues inside human body may have similar X-ray absorption rates and this phenomenon can considerably bias the estimations of the state-of-the-art patch-based methods.

The strength of our approach lies in the combination of multiple weak, but complementary visual cues, such as contour-based and region-based similarity measures all along the different steps of our registration/segmentation process. In addition, our superpixel-based framework that takes into account all the nonlinear and local variability of bone regions existing in the training dataset along with our label propagation fusion step, in the variation of information sense, combined with our outlier removal step, allows us to efficiently constrain the space of

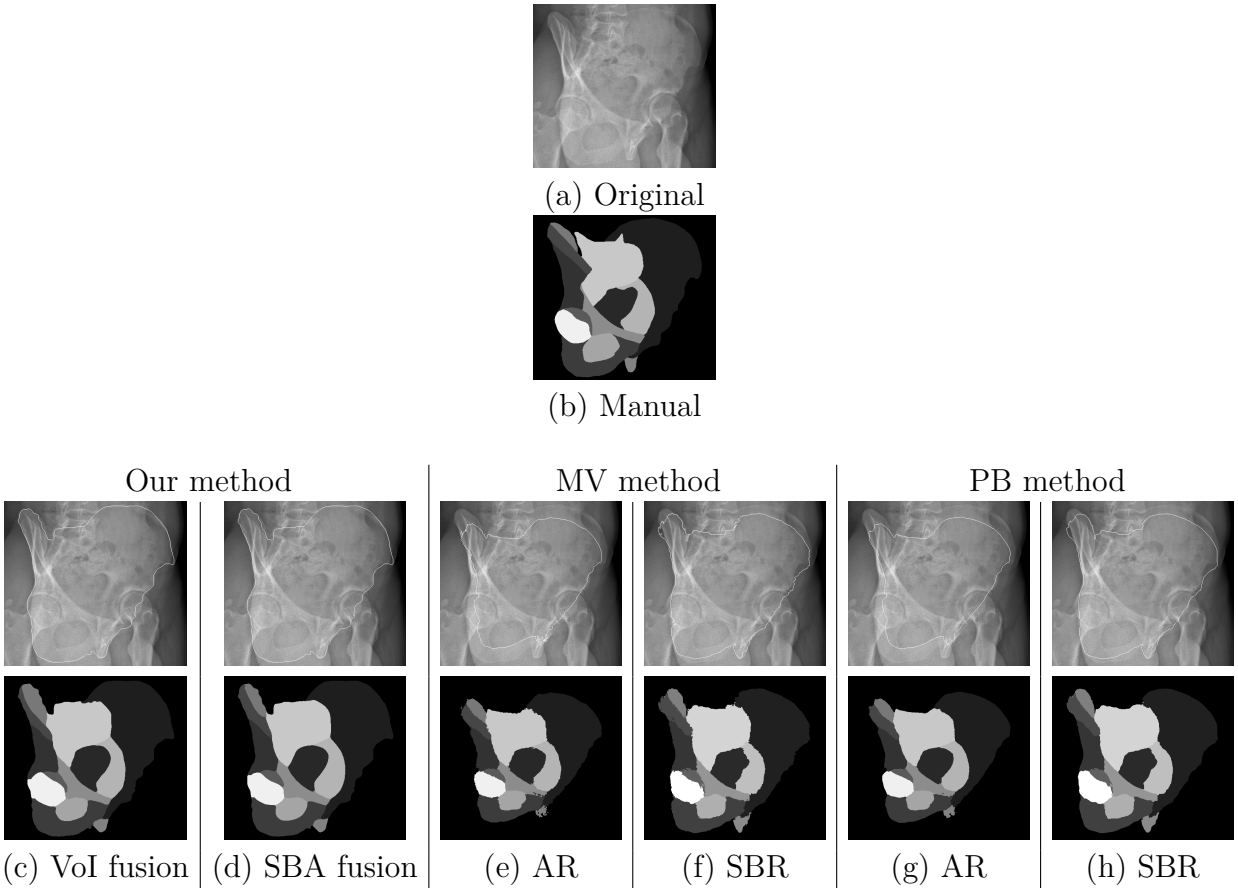
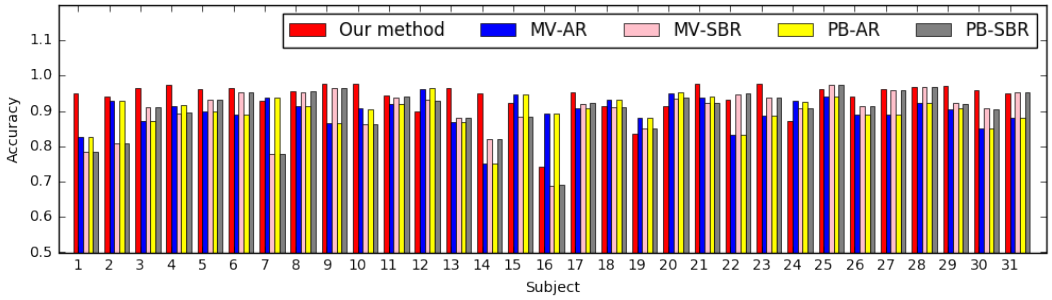
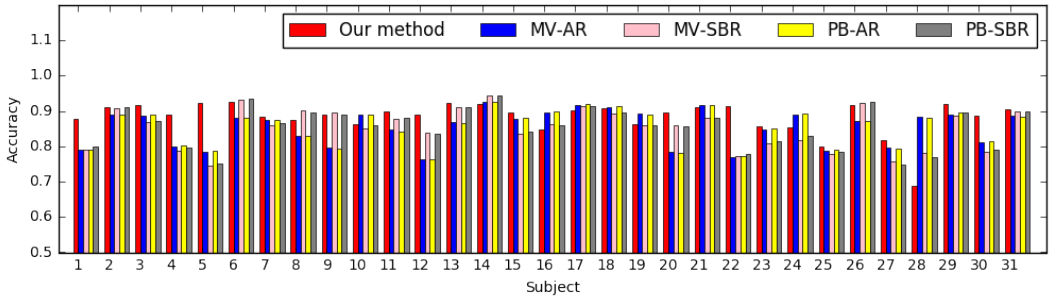


Fig. 17. Comparison of segmentation results of the pelvis. (a) original oblique X-ray radiographic image, (b) manual segmentation, third row: external pelvic contours on the corresponding X-ray image, fourth row: and resulting images.

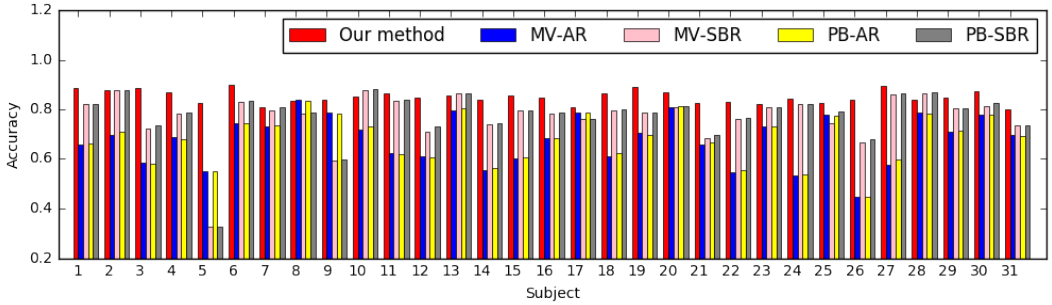
candidate segmentations, to finally propose a reliable segmentation map even in presence of low quality input X-ray images. This also makes our method original, robust and particularly appropriate for the segmentation of complex structures in X-ray imaging. To the best of our knowledge, no study has been done on the X-ray oblique image segmentation of the pelvis, talus, and patella. Unlike most existing multi-atlas-based segmentation approaches, the proposed method makes the following main contributions. First, the multi-atlas-based training dataset allows us to estimate, non-parametrically, a collection of superpixels which capture all the nonlinear and local variability of ROIs present in the multi-atlas dataset. This allows us both to increase the accuracy of the superpixel map pruning step while simplifying the optimization problem involved in this crucial step. Second, the contour-based similarity measure used in our first registration step and combined with the region-based similarity measure exploited in the final selection/filtering of the multi-atlas dataset allowed us to efficiently and adaptively single out the outlier segmentations, existing in the multi-atlas dataset, for each specific segmentation case. Third, the label propagation fusion step was herein performed in the variation of information sense and allowed us to refine the external



(a) Patella



(b) Talus



(c) Pelvis

Fig. 18. Accuracy of each subject in dataset. Each bin represents the similarity index of the resulting image comparing to the corresponding manual segmentation.

contour of the bone structure to be segmented, but also, and especially, to infer the internal region labels from the pre-selected segmentations of the atlas to the final segmentation result. In addition, the proposed approach is robust since it mainly relies on the external bone contour of the bone structure to be segmented, which is the most informative and reliable visual cues in an X-ray image, especially in the case of the lower limb bone structures (such as the pelvis, talus, and patella). Also, the average classification rate of our method was within the range of accuracy typically obtained in case of manual segmentations (w.r.t. intra inter/observer variability). The proposed segmentation approach has potential clinical

applications, one of them being three-dimensional reconstruction from biplanar images to assess skeletal morphology in the context of various pathological processes like osteoarthritis.

Deuxième article.

Unsupervised Registration of 3D Knee Implant Components to Biplanar X-ray Images

par

Dac Cong Tai Nguyen[1], Said Benameur[2], Max Mignotte[3] et Frédéric Lavoie[4]

([1]) DIRO, Université de Montréal et Eiffel Medtech Inc
([2]) Eiffel Medtech Inc
([3]) DIRO, Université de Montréal
([4]) Département de Chirurgie Orthopédique, CHUM et Eiffel Medtech Inc

Cet article a été soumis dans BMC Medical Imaging.

Mes contributions et le rôle des coauteurs

- DCTN a réalisé les travaux et rédigé le manuscrit.
- SB, MM et FL ont supervisé techniquement et biomédicalement ces travaux, et révisé le manuscrit.

Résumé. Le recalage des composants d'implants tridimensionnels (3D) du genou sur des images radiographiques fournit la position 3D des implants, ce qui aide à analyser l'alignement des composants après une arthroplastie totale du genou. Dans cet article, nous présentons un recalage automatique 3D à bidimensionnel (2D) utilisant des images radiographiques biplanaires basé sur une mesure de similarité hybride intégrant des informations sur la région et les contours. Plus précisément, cette mesure est définie comme une combinaison pondérée d'une similarité basée sur le champ de potentiel des contours, qui représente la relation entre les contours externes des projections des composantes et un champ de potentiel de contours estimé sur les deux images radiographiques, et une propriété de spécificité d'objet, qui est basé sur la distinction des régions à l'intérieur et à l'extérieur de l'objet. La précision de notre algorithme de recalage 3D/2D a été évaluée sur un échantillon de 64 composants (32 composants fémoraux et 32 composants tibiaux). Dans nos tests, nous avons obtenu une moyenne de l'erreur quadratique moyenne (RMSE) de 0,18 mm, ce qui est significativement inférieur à celui des deux méthodes de similarité simple, ce qui appuie notre hypothèse d'une meilleure stabilité et précision avec l'approche proposée. Notre méthode, qui fournit six paramètres de recalage (trois rotations et trois translations) précis sans nécessiter de marqueurs fiduciaux, permet d'effectuer les analyses importantes sur l'alignement en rotation des composants fémoraux et tibiaux sur un grand nombre de cas. De plus, cette méthode peut être étendue pour recaler d'autres implants ou os.
Mots clés : Recalage 3D/2D, images radiographiques, composants d'implants de genou, implants orthopédiques.

Abstract. Registration of three-dimensional (3D) knee implant components to radiographic images provides the 3D position of the implants which aids to analyze the component alignment after total knee arthroplasty. In this paper, we present an automatic 3D to two-dimensional (2D) registration using biplanar radiographic images based on a hybrid similarity measure integrating region and edge-based information. More precisely, this measure is herein defined as a weighted combination of an edge potential field-based similarity, which represents the relation between the external contours of the component projections and an edge potential field estimated on the two radiographic images, and an object specificity property, which is based on the distinction of the region-label inside and outside of the object. The accuracy of our 3D/2D registration algorithm was assessed on a sample of 64 components (32 femoral components and 32 tibial components). In our tests, we obtained an average of the root mean square error (RMSE) of 0.18 mm, which is significantly lower than that of both single similarity methods, supporting our hypothesis of better stability and accuracy with the proposed approach. Our method, which provides six accurate registration parameters (three rotations and three translations) without requiring any fiducial markers, makes it possible to perform the important analyses on the rotational alignment of the femoral and tibial components on a large number of cases. In addition, this method can be extended to register other implants or bones.
Keywords: 3D/2D registration, x-ray images, knee implant components, orthopaedic implants.

1. Background

Total knee arthroplasty (TKA) is an orthopaedic surgical procedure where the damaged articular surfaces of the knee joint are replaced with artificial implants. The implant consists of two metallic components that replace the bearing surfaces on the tibial plateau and femoral condyles, separated by a high molecular weight polyethylene insert. The component alignment after TKA has been proved as a significant factor in determining knee kinematics [**148, 164**], patellar tracking, and long-term clinical outcome [**101, 141**]. This alignment is currently evaluated in 2D X-ray images [**140, 139**]. However, position of the X-ray source, orientation of the subject's pelvis and lower extremity may have an effect on measurements obtained from 2D radiographs. A 3D analyses of component positions after TKA will possibly not only to increase the accuracy of measurements, but also to lead to new works on TKA, or to improve implant designs which increase their life span because abnormal knee kinematics might cause premature failure of the implant [**88, 29**].

A 3D lower extremity alignment assessment system has been created to evaluate 3D alignment by manually matching 3D bone and component projections with frontal and oblique X-ray images of the entire lower extremity [**141, 142**]. But this system is time-consuming and has low accuracy of position estimation [**85**]. 3D alignment information after TKA can also be obtained from magnetic resonance imaging (MRI) and computed tomography (CT) scan [**100, 102**] but is costly and involves significant radiation exposure in the case of CT scans. If serial follow-up evaluations are needed, the cost and the issue of radiation will greatly increase after sequential MRI or CT scans. Another approach based on Roentgen stereophotogrammetry has been developed [**161**]. Roentgen stereophotogrammetric analysis (RSA) is a highly accurate technique for 3D micromotion evaluation of orthopaedic implants but it is limited by the need to surgically insert numerous tantalum beads into the bones with a special instrument.

In our case, a 3D/2D registration based method seems to be an adequate and suitable solution. 3D/2D registration methods have been used in many medical fields, mostly in image-guided therapy, such as cancer diagnosis and therapy [**172, 31**], radiosurgery [**47**], interventional radiology [**97**], and a variety of therapies in surgery [**128, 65, 152**]. In this context, these techniques align 3D implant components to 2D X-ray images to determine their 3D information (positions and orientations). These registration approaches can be classified into two groups based on the number of images used. Most of them use fluoroscopic images for 2D X-ray image.

The first group uses a single X-ray image and biplanar X-ray images are used in the second group. Previous approaches used a direct image to image similarity measure [**96**] or contour-based 3D/2D registration [**173, 67**] to estimate the pose of knee implant. Such technique can provide clinically sufficient accuracy only for five degrees of freedom (DOF)

(three rotations and two translations parallel to image), as the DOF related to the translation perpendicular to image (depth position) is quite challenging. Yamazaki *et al.* [**173**] improved the depth position estimation by optimizing it independently of the five other DOFs, using an approximate evaluation curve of depth position prepared after initial registration. Although depth position was improved, it wasn't judged to be sufficiently accurate. The RMSE, average errors and standard deviation of depth position in this technique were ten times higher than those of two other translations. Another approach to increase the accuracy of depth position estimation was based on the fluoroscopic imaging property that the closer the object is to the source, the larger the image at the image intensifier is produced. Hossain *et al.* [**71**] determined the scale change due to the depth translation by using a calibration box to estimate the depth position. This approach gave accurate results but required an extra step to compute the scale change in the depth translation of the fluoroscopy unit. The second limitation is due to the high degree of shape symmetry of the components, particularly the tibial component. A symmetrical pose might be obtained instead of the true one because the representation of both solutions on the X-ray image is very similar (see Fig. 19). In [**67**], the

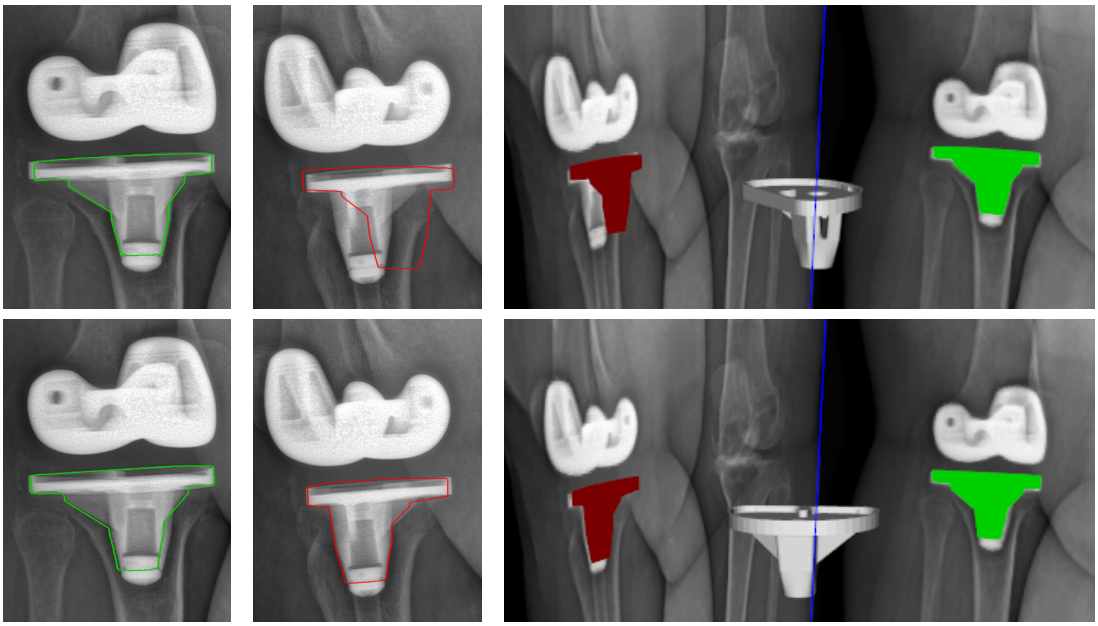


Fig. 19. Example of tibial component projection on two X-ray images. From top to bottom : symmetrical and true pose. (a), (b) component external contours projected on 135-degree image and 45-degree image, and (c) 3D view of component projections.

authors try to solve this problem by simultaneously estimating two symmetrical poses, but the algorithm still might not converge toward the true pose. By using biplanar 2D images, these methods can measure all six DOFs with a sufficient accuracy (see Fig. 19) and avoid the symmetrical issue. Kim *et al.* [**83**] optimized the normalized correlation coefficient (NCC)

between the dual X-ray images and two corresponding virtually projected component images to obtain six DOFs. George *et al.* have shown that using biplanar images provides a highly accurate estimation of six DOFs [**64**]. However, biplanar fluoroscopic images at the same exact position are not easy to obtain due to the moving pictures and the complexity of the system. In addition, they involve a potential risk of radiation to the patient. Radiography is an adequate choice as it is the lowest in term of cost, complexity, and risk of radiation compared to fluoroscopy, MRI, or CT scan.

In this paper, we propose a simple method to align the components of a 3D knee implant to biplanar oblique X-ray images. This method does not require fiducial markers and/or intraoperative X-ray image segmentation. It uses a hybrid (relying on both region and edge-based information) similarity measure both combining the contour similarity between the external contours of the component projections and an edge potential field estimated on the two radiographic images [**75**] and a region label similarity measure term (which is based on the distinction of the region label inside and outside of the object [**46**]). Then a stochastic Exploration Selection (ES) algorithm is used to estimate the six DOFs of implant position.

2. Proposed Approach

2.1. Image Pre-processing

Contour detection. Due to the metallic material, the implant components appear much whiter than the neighboring bones and soft tissues. A pre-processing process is performed on each image to enhance the visibility of the component contours which, in fact, constitute the most important and reliable low-level visual cue in each radiographic image. First, a histogram equalization technique increases the global contrast. Second, a median filter and non-local means denoising [**28**] algorithm are used to remove the noise from the images. The non-local means denoising method replaces a pixel with a weighted average of pixels having a similar neighborhood. More precisely, for each pixel, it first searches in a large search window (centered on the pixel to be denoised) for all the neighborhoods of pixels that most closely resemble (with a least squares (LSQ) similarity measure) the neighborhood of the pixel to be filtered. Then, a weighted average (based on the previous LSQ similarity measure) of all these central pixels (of all these neighborhoods) allows to estimate the denoised greyscale value of the pixel. Finally, the edges are detected by using a Canny edge filter [**30**] (see Fig. 20).

Label detection. The next part of the pre-processing step is the region-label extraction. By using the Simple Linear Iterative Clustering (SLIC) algorithm [**2**], the input images are segmented into K labels (*i.e.* superpixels). This algorithm is actually simple and has a low computational cost. SLIC performs a clustering of pixels in a five-dimensional (5D) space

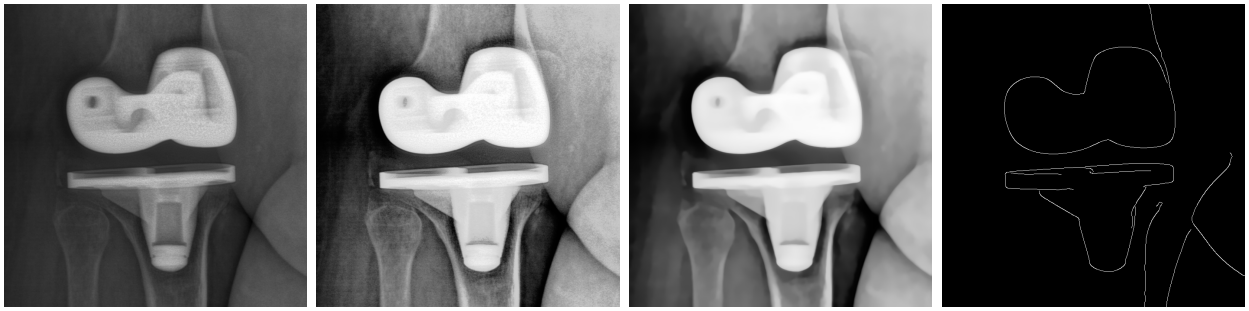


Fig. 20. Example of contour-based pre-processing step. From left to right : input X-ray image, histogram equalization image enhancement, denoised image, pre-processed image.

based on their color similarity and proximity in the image. The 5D space is defined by the L, a, b values of the CIELAB color space and x, y coordinates of the pixels. Due to the fact that the distance between two colors in the CIELAB space is different from the spatial distance in the xy plane, it is not possible to simply use the Euclidean distance. In order to cluster pixels in this 5D space, a new distance measure based on Euclidean distance, with normalization of the spatial distances, was introduced:

$$d_{Lab} = \sqrt{(L_u - L_v)^2 + (a_u - a_v)^2 + (b_u - b_v)^2} \tag{2.1}$$

$$d_{xy} = \sqrt{(x_u - x_v)^2 + (y_u - y_v)^2} \tag{2.2}$$

$$D = d_{Lab} + \frac{m}{S} d_{xy} \tag{2.3}$$

where D is the sum of the *Lab* distance and the *xy plane* distance normalized by the grid interval $S = \sqrt{\frac{N}{K}}$, N is the number of pixels in image and m is a variable to controlling the compactness of a superpixel.

The algorithm begins by initializing K cluster centers. Each pixel in the image is associated with the nearest cluster center whose search area overlaps this pixel. After all the pixels are associated, new centers are computed as the average 5D vector of all the pixels belonging to the cluster. The assignment of each pixel to the nearest cluster center and the re-computation of the new cluster center process are iteratively repeated until convergence. At the end of this process, a few remaining pixels are enforced to connect to the largest neighboring cluster (see Fig. 21).

2.2. 3D/2D Registration Procedure

Once the edges and the region-based labels are extracted from the image by our pre-processing step (see "Image Pre-processing" section), a rigid registration is performed to align the component to biplanar images. To this end, we combine the object specificity property [**46**] and the similarity between the external contours of the component projections and an edge potential field estimated on the two radiographic images.

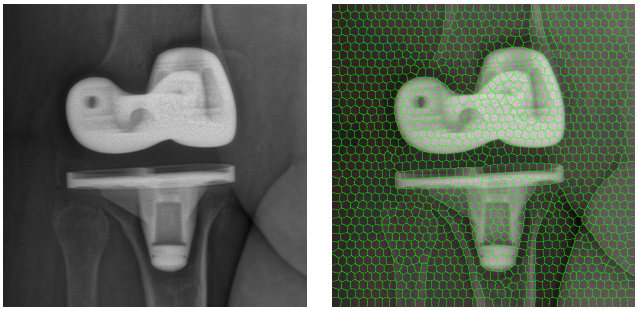


Fig. 21. Example of label-based pre-processing step : input X-ray image (left) and pre-processed image (right).

2.2.1. Similarity measurement.

Edge potential field-based similarity. An edge potential field-based similarity measure is defined in order to evaluate the concordance or the similarity degree between the external contours of the component projections on biplanar X-ray images and an edge potential field, estimated from the previously detected edges of the two views. Concretely, this edge potential field attracts the component and aligns it to the edge from the input image by giving a similarity measure all the greater as the edges of the projected contours of the component coincide well with the edges existing in the two views. The edge potential field Φ of each view is computed on the pre-processed image and is defined (as in [**75**]) as:

$$\Phi(x,y) = \exp(-\rho\sqrt{\delta_x^2 + \delta_y^2}) \tag{2.4}$$

where (δ_x,δ_y) is the displacement to the nearest edge point in the image, and ρ is a smoothing factor which controls the degree of smoothness of the potential field. Finally, a directional component is added to produce a cost function measuring the correspondence between the projected contours of the component and the edges in the two views:

$$\begin{aligned} E = &-\frac{1}{n_1}\sum_{n_1}(\Phi_1(x,y)|\cos(\alpha_1(x,y))|) \\ &-\frac{1}{n_2}\sum_{n_2}(\Phi_2(x,y)|\cos(\alpha_2(x,y))|) \end{aligned} \tag{2.5}$$

where $\alpha(x,y)$ is the angle between the tangent direction of the external contours at (x,y) and the tangent of the nearest edge, and n_1,n_2 are the number of pixels on external contours of the component of each view (see Fig. 22).

Object specificity property. Object specificity property is based on the following hypothesis : *the labels inside and outside the object are distinct.* This property is fulfilled (thus involving a minimal cost or error function) whenever the labels inside an object are specific to that object, *i.e.* the labels inside an object do not occur outside that object and vice versa. It is

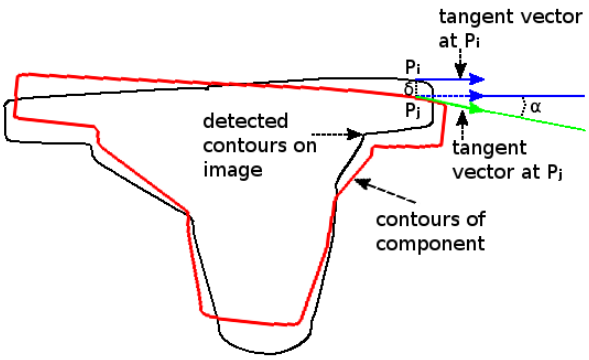


Fig. 22. Directional component used in equation (2.5).

defined by:

$$\vartheta = \sum_{h=1}^{K} p_h |\{s : z_s = l_h, s \notin c^{int}\}| \tag{2.6}$$

where $\Lambda = \{l_1,...,l_K\}$ is the set of K labels in image, $z_s \in \Lambda$, p_h is the proportion of the pixels with label l_h in the interior c^{int} of component, and the last factor is the number of pixels belonging to label l_h in the exterior of component. To reduce execution time, the last factor can be written as $|\{s : z_s = l_h\}| - |\{s : z_s = l_h, s \in c^{int}\}|$, so $|\{s : z_s = l_h\}|$ can be precomputed, only p_h is computed dynamically. By normalizing this property, we obtain a cost (or error) function (to be minimized):

$$V = \frac{\vartheta_1}{N_1} + \frac{\vartheta_2}{N_2} \tag{2.7}$$

where N_1,N_2 are the number of pixels of component related to each view.

2.2.2. Six registration parameters optimization. Finally, our registration approach is formulated as the minimization of the following cost function:

$$C(\theta) = E(\theta) + \beta V(\theta) \tag{2.8}$$

where θ is a set of six registration parameters and β is a weighting factor between respectively the edge-based and region-based energy terms of our energy-based registration model. In order to minimize this complex non-convex cost function, we resort to ES algorithm, a stochastic and efficient optimization algorithm, already proposed in [**60**] and especially well-suited for this type of function to be optimized.

In fact, the ES algorithm belongs to the class of evolutionary algorithms. This class of algorithm inspires the natural evolution to solve hard problems. Suppose that a problem is a natural environment which encompasses a population of individuals. Each individual represents a possible solution to the problem. A fitness function is used to measure the degree of adaptation of each individual (*i.e.* potential solution) to its environment (*i.e.*

problem). Like evolution in nature, these algorithms produce progressively better solution to the problem. This class of algorithm has been successfully used in medical imaging [**127, 115, 18**] or for detection and accurate localization of shapes in traditional images [**46**]. More formally, the ES algorithm can be described in two steps: exploration and selection steps (more details are given in [**60**]). The first step implicates a probabilistic operator to attempt a random search on a graph of each individual which is considered as a potential solution. And the second step creates interaction and selection between individuals. This process is run until a stopping criterion has been met (cf. Algorithm 5).

Algorithm 5 ES Algorithm

$C(.)$	A k-variable function to be minimized
$\mathcal{N}(a)$	The neighborhood of an element a
k	A number of intervals
D	$(= k/r)$ The diameter of the exploration graph
θ	A set of registration parameters
$\hat{\theta}$	The optimal element θ_i of θ
p	The probability of exploration
t	The iteration step
n	The size of the population (greater than D)
r	A real number $\in$ [0,1] called the radius of exploration

1. Initialization
Random initialization of 6 registration parameters
$t \leftarrow 2$

2. Exploration/Selection
while *a stopping criterion is not met* ***do***
(1) *Compute* $\hat{\theta}$; $\hat{\theta} \leftarrow \arg\min_{\theta_i \in \theta} C(\theta_i)$
(2) *Draw* f *according to the binomial law* $b(n,p)$
- *For* $i \leq f$, *replace* θ_i *by* $\gamma_i \in \mathcal{N}(\theta_i) \setminus \{\hat{\theta}\}$ *according to the uniform distribution (Exploration step)*
- *For* $i > f$, *replace* θ_i *by* $\hat{\theta}$ *(Selection step)*

(3) $t \leftarrow t+1$ *and* $p \leftarrow t^{-1/D}$
end while

3. Experimental Results

To evaluate the accuracy of our 3D/2D registration algorithm, ground truth models were created. A ground truth model is the rendering of the 3D implant components on a real X-ray image (see Fig. 23). The ground truth image retains the properties of the X-ray imaging such as variety of imaging noise originated from several components of the system (X-ray source, CCD camera, controller circuits, etc.), and the patient bone structures overlap with other bones or dense soft tissues (cartilage, meniscus, and fascia). A random transformation (rotation and translation) was applied to the ground truth 3D components. Then

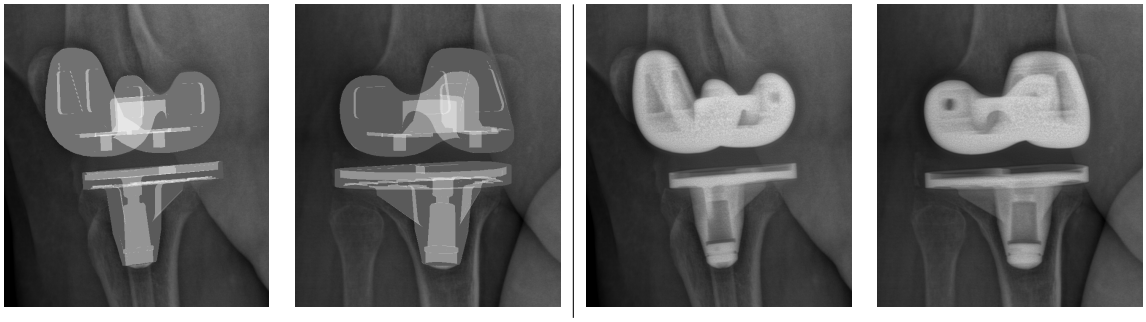


Fig. 23. Example of biplanar ground truth images (left) and biplanar real radiographic images (right).

we compared the registered transformed components with the ground truth components. This comparison was evaluated on 64 randomly transformed components (32 femoral and 32 tibial components). In this comparison, we also performed the registrations with different similarity measurements. Afterward, these results were analyzed by using a statistical Z-test [1] between hybrid similarity and single similarity approaches where the null hypothesis was that there is no difference between two approaches, and a statistical significance threshold set at $p \leq 0.001$. Then the p-value was calculated.

Table 5 shows the average RMSE [2] for the random transformed components before registration. Table 6 shows the average RMSE for each component and for all components by using potential field similarity, object specificity similarity and our hybrid similarity in our accuracy test, as well as the comparison between the different approaches. Table 7 shows the average errors (AE) and the standard deviations (SD) by using single similarities and hybrid similarity, and the comparison between each DOF. Fig. 24 shows the RMSE of each transformed components. Finally Fig. 25 and 26 show an example of 3D/2D registration on ground truth images and real radiographic images, respectively.

	Initial RMSE
Femoral components	12.055
Tibial components	11.481
All components	11.768

Tableau 5. Average of RMSE (mm) for the random transformed components before registration.

In our experiments, we set the size of the population to 20 and the number of iterations to 800. For the tibial components, the average RMSE were 2.66 mm, 0.28 mm and 0.18

[1] Z-score$=\frac{\mu_2-\mu_1}{\sqrt{\sigma_1^2/n}}$, where μ_2 is the average error of the single similarity approach, μ_1 and σ_1^2 are the average error and the variance of the hybrid similarity approach, and n is the number of components.

[2] $RMSE = \sqrt{\frac{\sum_{i=1}^{n} dist_i^2}{n}}$, where $dist_i$ is the distance between a vertex of the transformed component and its correspondence in the ground truth component, and n is the number of vertices.

	Hybrid	**Potential field**	**Object specificity**
Femoral components	0.177	0.525	0.269
Tibial components	0.183	2.658	0.28
All components	0.18	1.592	0.275
p-value$_{Hybrid/...}$		< 0.0001	< 0.0001

Tableau 6. Accuracy test (average of RMSE (mm) and p-value).

		Rotation (degrees)			**Translation** (mm)		
		X	Y	Z	X	Y	Z
Femoral components	Hybrid	1.047 ± 0.771	0.767 ± 0.591	1.062 ± 0.672	0.046 ± 0.03	0.047 ± 0.031	0.086 ± 0.043
	Potential field	1.065 ± 0.87	0.822 ± 0.68	0.879 ± 0.771	0.187 ± 0.146	0.161 ± 0.115	0.142 ± 0.127
	Object specificity	1.02 ± 0.767	0.869 ± 0.674	0.977 ± 0.722	0.075 ± 0.049	0.065 ± 0.052	0.101 ± 0.07
Tibial components	Hybrid	0.727 ± 0.567	0.713 ± 0.473	0.878 ± 0.578	0.085 ± 0.046	0.075 ± 0.039	0.046 ± 0.035
	Potential field	2.635 ± 1.9	4.854 ± 3.492	3.241 ± 1.765	0.829 ± 0.368	0.777 ± 0.67	0.561 ± 0.554
	Object specificity	0.641 ± 0.523	0.776 ± 0.595	0.869 ± 0.61	0.073 ± 0.062	0.109 ± 0.068	0.067 ± 0.05
All components	Hybrid	0.887 ± 0.695	0.74 ± 0.536	0.97 ± 0.634	0.065 ± 0.044	0.061 ± 0.038	0.066 ± 0.044
	Potential field	1.85 ± 1.673	2.838 ± 3.224	2.06 ± 1.803	0.508 ± 0.426	0.469 ± 0.571	0.351 ± 0.453
	p-value$_{Hybrid/PotentialField}$	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
	Object specificity	0.83 ± 0.684	0.822 ± 0.637	0.923 ± 0.67	0.074 ± 0.056	0.087 ± 0.064	0.084 ± 0.063
	p-value$_{Hybrid/ObjectSpecificity}$	0.5093	0.2225	0.5552	0.0989	< 0.0001	0.0012

Tableau 7. Average errors ± standard deviations and p-values of six DOFs.

mm by using the edge potential field-based similarity measure, the object specificity similarity measure and our hybrid similarity measure, respectively. The average RMSE of the femoral components in these three similarity measures were 0.53 mm, 0.27 mm and 0.18 mm, respectively. Finally, for all components, the average RMSE were 1.59 mm, 0.28 mm and 0.18 mm, respectively, and the p-values between hybrid similarity and single similarity approaches were inferior than 0.0001. A complete evaluation took, on average, approximately 245 seconds on a 64-bit desktop PC (Ubuntu 16.04 LTS, 1.30GHz Core i7 CPU and a graphic card with Intel, 16GB RAM). Note that we didn't use any libraries in programming in C++.

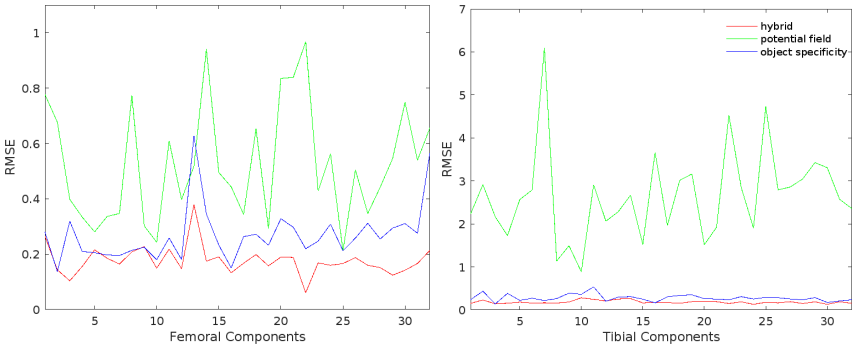
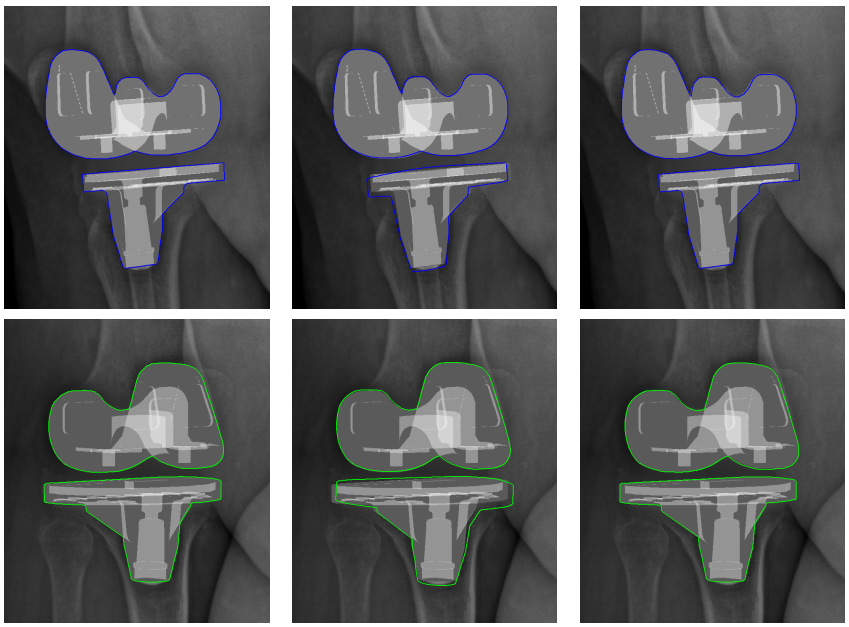


Fig. 24. RMSE of each transformed components.

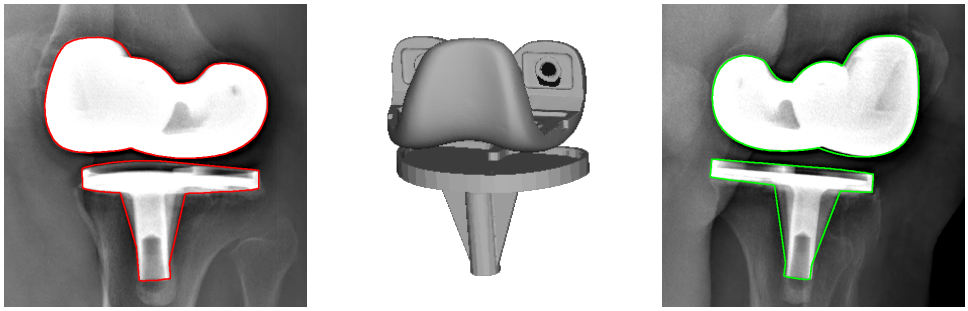


(a) (b) (c)

Fig. 25. Example of 3D/2D registration result on ground truth images: (a) hybrid similarity, (b) potential field similarity, and (c) object specificity similarity, first row : 45-degree image and second row : 135-degree image.

4. Discussion

Our tests showed that the potential field similarity is sensitive. Its average RMSE of the tibial components was a lot higher than those of the femoral components because there are more artifacts around the tibial component than the femoral component. Tibial component contours are attracted by the potential field of the tibial and fibular contours. The object specificity similarity was more stable and accurate than the potential field similarity. But the combination of both edge and region-based similarity measures gave the best result in terms of stability and accuracy, as shown by the statistically significant difference that was found between hybrid similarity and single similarity approaches. The difference of their



(a) (b) (c)

Fig. 26. Example of 3D/2D registration result on real radiographic images. (a), (c) component external contours projected on 45-degree image and 135-degree image, and (b) 3D view.

average RMSE of each component was less than 0.006 mm. The advantage of label-based similarity is its stability. In addition, this term, based on the region process and a numerical integration is inherently robust to noise. However, this measure alone is not precise because the number of pixels in the border labels is fewer than the number of pixels inside the label. On the contrary, the edge-based similarity measure is accurate while also being more sensitive to noise or other artifacts existing in the images. That's why the combination of the stability and the robustness to noise of the label-based similarity term with the accuracy of the edge-based similarity measure, provides a robust unsupervised registration method.

Compared to the NCC-based method in [**83**], the experiments showed that their results are very slightly better than that of our method (see Table 8). The differences were less than 0.04 mm. Note that the tests of the NCC-based method used fluoroscopic images of sawbones and didn't involve the rotations.

	Our method	**NCC-based method**
Femoral components	0.177	0.141
Tibial components	0.183	0.15
All components	0.18	0.146

Tableau 8. Results (average of RMSE (mm)) of our method versus the NCC-based method in [**83**].

Unlike other works that use fluoroscopic images, our method uses biplanar radiographic images which are the advantageous in terms of cost, complexity, and risk of radiation and provide six registration parameters with a sufficient accuracy without the need for additional fiducial markers. Our method is also robust to image noise and occlusions, as demonstrated by the AE and the SD of each parameter in our tests. The AE of translation parameters were around 0.06 mm and their SD were less than 0.05 mm. For the rotation parameters, the

AE were less than 1 degree, and the SD were around 0.6 degree. In addition, this method can be extended to register other implants or bones to biplanar oblique or frontal/lateral X-ray images (see Fig. 27, 28).

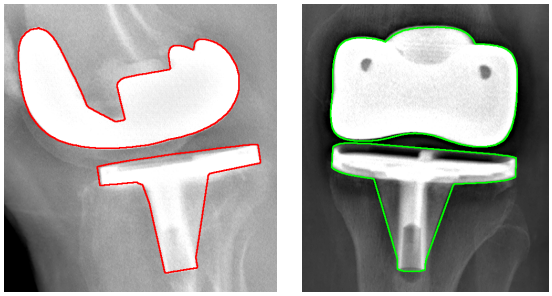


Fig. 27. Example of our registration for femoral and tibial components on real radiographic images : lateral image (left) and frontal image (right).

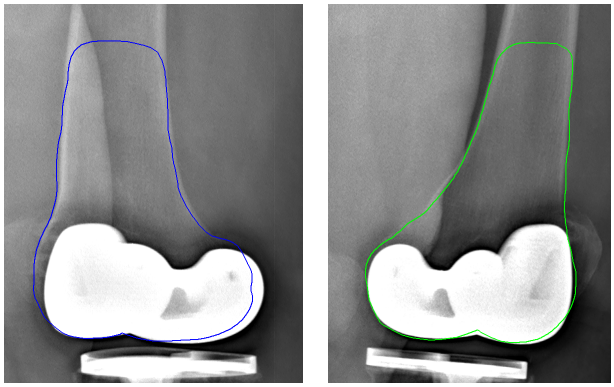


Fig. 28. Example of our registration for the distal femur on real radiographic images: 45-degree image (left) and 135-degree image (right).

The proposed method, based on the ES optimizer, is slightly time-consuming but it is easily parallelizable and thus remains especially well-suited for the next-generation GPU or massively parallel computers and multi-core processors.

Based on the result of our TKA component registration, the rotational alignment of the femoral and tibial components can be studied by measuring and analyzing both component position and orientation. For example, the external rotation angles of the implants can be measured. These rotations are important in patello-femoral tracking because inappropriate rotation of the femoral component may cause flexion imbalance and patellofemoral problems [**83**]. The combined rotational alignment change after TKA can also be measured and the different influence of symmetric and asymmetric tibial component designs on the combined rotational alignment can be compared [**132**]. Our accurate registration method makes it possible to perform these important analyses on a large number of cases.

5. Conclusion

We have presented an unsupervised registration of 3D knee implant components to biplanar X-ray images. This method uses a hybrid similarity measure by combining the object specificity property and the similarity between the external contours of the component projections and an edge potential field (related to the edges) estimated on the two radiographic images. A stochastic optimizer (ES) algorithm is then used to efficiently estimate the six DOFs of implant position. Our method can avoid symmetrical solution and provides six registration parameters with a good accuracy. Moreover, it does not require any fiducial markers. The proposed 3D/2D registration approach has the potential to increase the effectiveness of computer-aided clinical analysis, namely relative angle analysis which is important to predict not only the function but also the stability and survival of TKA implants [**21**].

Troisième article.

3D Biplanar Reconstruction Of Lower Limbs Using Nonlinear Statistical Models

par

Dac Cong Tai Nguyen[1], Said Benameur[2], Max Mignotte[3] et Frédéric Lavoie[4]

([1]) DIRO, Université de Montréal et Eiffel Medtech Inc
([2]) Eiffel Medtech Inc
([3]) DIRO, Université de Montréal
([4]) Département de Chirurgie Orthopédique, CHUM et Eiffel Medtech Inc

Cet article a été soumis dans Elsevier Medical Image Analysis.

Mes contributions et le rôle des coauteurs

- DCTN a réalisé les travaux et rédigé le manuscrit.
- SB, MM et FL ont supervisé techniquement et biomédicalement ces travaux, et révisé le manuscrit.

Résumé. La reconstruction tridimensionnelle (3D) des membres inférieurs joue un rôle important dans la planification chirurgicale, la chirurgie assistée par ordinateur et les applications biomécaniques. L'utilisation de modalités d'imagerie 3D telles que la tomodensitométrie (CT-Scan) et l'imagerie par résonance magnétique (MRI) présente des limites telles que les coûts et la dose de rayonnement élevée. Par conséquent, les méthodes de reconstruction 3D à partir d'images radiographiques biplanaires sont des méthodes alternatives afin d'obtenir des modèles 3D précis avec un coût et une dose de rayonnement moins élevés. Dans cet article, nous présentons une nouvelle approche de reconstruction 3D non supervisée pour la rotule, l'astragale et le bassin. Notre approche utilise des images radiographiques conventionnelles biplanaires calibrées (obliques à 45 et 135 degrés) et une information *a priori* sur la structure géométrique de ces os complexes. Un algorithme de réduction de dimensionnalité non linéaire basé sur un positionnement multidimensionnel (MDS) est appliqué pour exploiter ces informations géométriques *a priori*. Il représente les déformations pertinentes existant dans la base d'apprentissage. La méthode proposée repose sur une vraisemblance hybride utilisant des régions et des contours. La notion basée sur les contours représente la relation entre les contours externes des projections des os et un champ de potentiel de contours estimés sur les images radiographiques. La notion basée sur les régions est le taux de non recouvrement entre les régions d'intérêt (ROIs) des os segmentées et celles projetées. Ensuite, notre modèle de reconstruction 3D automatique aboutit finalement à minimiser stochastiquement une fonction d'énergie permettant d'estimer les paramètres de déformation de la forme osseuse. Cette méthode de reconstruction 3D a été testée avec succès sur 13 couples d'images radiographiques biplanaires, donnant des résultats très prometteurs.
Mots clés : Reconstruction 3D, images radiographiques, radiographies biplanaires, modèles statistiques non linéaires, imagerie médicale.

Abstract. Three-dimensional (3D) reconstruction of lower limbs is of great interest in surgical planning, computer assisted surgery, and biomechanical applications. The use of 3D imaging modalities such as computed tomography (CT) scan and magnetic resonance imaging (MRI) have limitations such as costs and high radiation dose. Therefore, 3D reconstruction methods from biplanar X-ray images represent an alternative to achieve accurate 3D models with lower cost and radiation dose. In this paper, we present a new unsupervised 3D reconstruction approach for the patella, talus, and pelvis. Our approach uses calibrated biplanar (45- and 135-degree oblique) radiographic images, and a prior information on the geometrical structure of these complex bones. A multidimensional scaling (MDS) based nonlinear dimensionality reduction algorithm is applied to exploit this prior geometrical information. It represents relevant deformations existing in the training set. The proposed method reposes upon a hybrid likelihood using regions and contours. The edge-based notion represents the relation between the external contours of the bone projections and an edge potential field estimated on the radiographic images. Region-based notion is the non-overlapping ratio between segmented and projected bone regions of interest (ROIs). Then, our automatic 3D reconstruction model finally results in stochastically minimizing an energy function allowing to estimate the deformation parameters of the bone shape. This 3D reconstruction method has been successfully tested on 13 biplanar radiographic image pairs, yielding very promising results.
Keywords: 3D reconstruction, X-ray images, biplanar radiographies, nonlinear statistical models, medical imaging.

1. Introduction

Three-dimensional (3D) reconstruction of bones is an important step for various clinical applications. It may be used in clinical studies for diagnosis and treatment of pathologies, especially related to the bony structure of the human body and precise implant selection, among others. 3D models are used more extensively for operation planning and morphometric studies. Applying personalized 3D precise models by surgeons has great potential impact on the accuracy and reliability in orthopaedics. X-ray images are often used due to their wide availability, lower price, and levels of ionizing radiation compared to computed tomography (CT) scan and magnetic resonance imaging (MRI). However, since X-ray images provide only 2D information, some prior knowledge must be incorporated in order to extract the missing dimension.

Previous approaches [**13, 156, 22, 124, 84**] use *a prior* knowledge on the geometrical structure of the object to be reconstructed. Nevertheless, in these methods, the geometric *a priori* constraint does not model the set of admissible deformations of the anatomical structure to be reconstructed. Consequently, the 3D shape estimation does not necessarily correspond to the reality.

Methods using statistical *a priori* knowledge of the geometric shapes of the objects of interest lead to better constrain the reconstruction problem [**178, 176, 12, 34, 52, 59**].

These methods use statistical shape models or statistical shape and intensity models for reconstructing bones from X-ray images. However, the optimization of the deformable model parameters might be slow and needs a good initialization point to avoid local maxima [**131, 82**]. In [**18, 17**], Benameur *et al.* used the contours of anatomical shapes detected in the two radiographic images for the 3D reconstruction of the anatomical object considered. However, segmentation or edge detection algorithms are very sensitive to noise and this often results in noisy contours, which are difficult to use in the case of these images.

In [**81**], Kasten *et al.* used deep learning for 3D reconstruction of bones from two bi-planar X-ray images. However, implementing 3D reconstruction from two or more 2D images using a deep learning approach remains a challenging task, due to the difficulty of representing a dimensional enlargement in multi-view settings with standard differentiable layers. Moreover, due to the transparent nature of X-ray images, matching surface points between multi-views for dense reconstruction is extremely challenging compared to the standard multi-view setting. A survey on deep learning in medical image reconstruction is given in [**4**]. In addition, deep learning methods require a lot of computing power, are expensive to set up, make decisions that are difficult or not at all understandable, and require a large database. [**78**] gives a comparative survey on 3D reconstruction of medical modalities based on various approaches.

The 3D reconstruction approach of bones that we propose will be obtained thanks to the calibrated 45- and 135-degree oblique radiographic images[3] and a general *a priori* knowledge of the geometric structure of each bone. This knowledge comes from a considerable learning base of surface models. This *a priori* knowledge is obtained by applying the variant of Multidimensional Scaling (MDS) algorithm presented in [**32, 25**] by choosing as the metric the most adequate distance to our learning base. MDS allows to find a small-dimensional subspace which preserves the chosen metric in the original space. This step of nonlinear dimensionality reduction will allow us to define a mesh of deformation in a reduced dimensional space in which each point of this mesh will, in fact, correspond to a surface model of our learning base. As for the Principal component analysis (PCA) in [**18**], it will allow us to define a concise 3D parametric reconstruction model in which the set of statistically admissible (in our case, possibly non-linear) deformations and learned on our learning base, will be summarized by the values of a reduced vector of parameters.

From this mesh of non-linear deformations which summarizes all the statistically admissible surface structures, the proposed reconstruction method will then consist in adjusting the projections (45-degree and 135-degree oblique) of a surface model of bone (contained in this deformation mesh and by interpolation) with the contours and the regions of segmented images, containing the corresponding bone. This approach is based on a likelihood using

[3]Due to the physical link between the detector source assemblies, the position in space of the sensors and X-ray sources are well known: the radiographic environment is therefore pre-calibrated.

a prior detection of the contours (oblique projections of the 3D model) as in [**18**] but also on a global constraint using the notion of specificity. This property is based on the fact that labels (of classes), given by a prior textural over-segmentation of the image, inside and outside the shape, are distinct (or specific in a neighborhood of the object). Thanks to this likelihood, using both edge and region information, our Bayesian reconstruction model has the property of being particularly robust to noise. Finally, the 3D reconstruction problem is then seen as a simple problem of estimating the deformation parameters of this 3D surface model or, equivalently, as a problem of minimizing a cost function. This minimization will be efficiently carried out by a stochastic algorithm.

The remainder of this paper is divided into the following sections : Section 2 describes our approach with the pre-processing step made on each X-ray image to enhance the contours and extract the regions, the statistical deformable model, energy function, and the optimization algorithm. We show experimental results in Section 3, discuss and conclude in Section 4.

2. Proposed Model

Our 3D reconstruction approach uses two calibrated biplanar oblique 45-degree and 135-degree radiographic images and *a prior* global knowledge of the geometric structure of each bone. We assume the oblique orthogonal radiologic projections will allow a better identification of the various components of the bone.

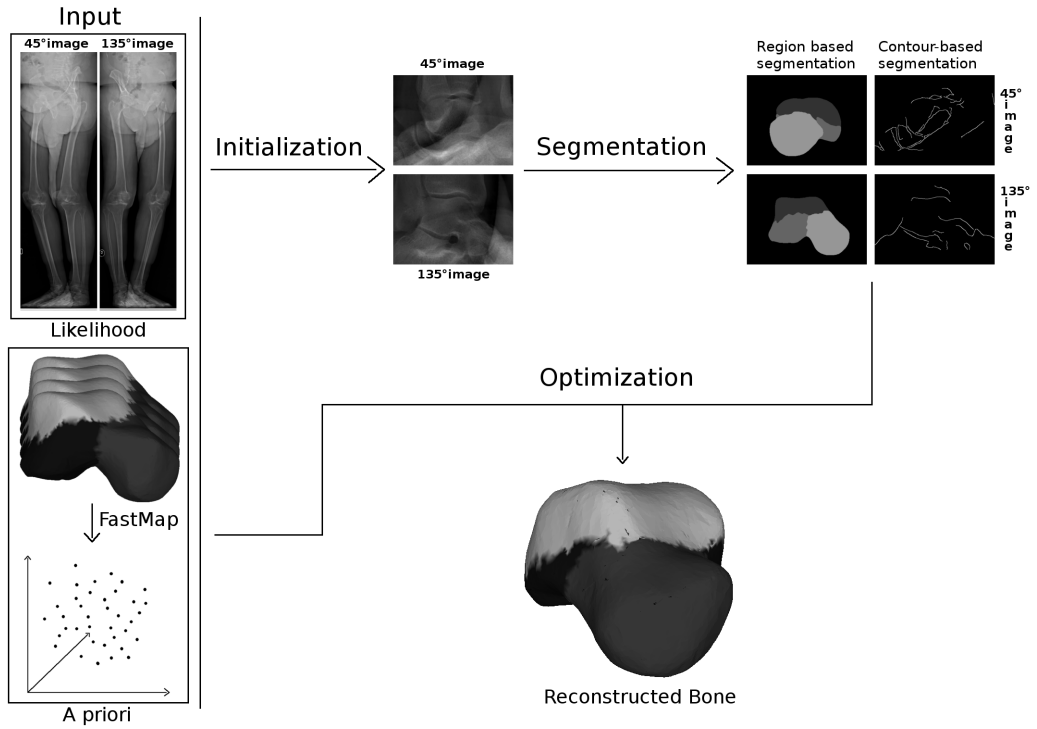
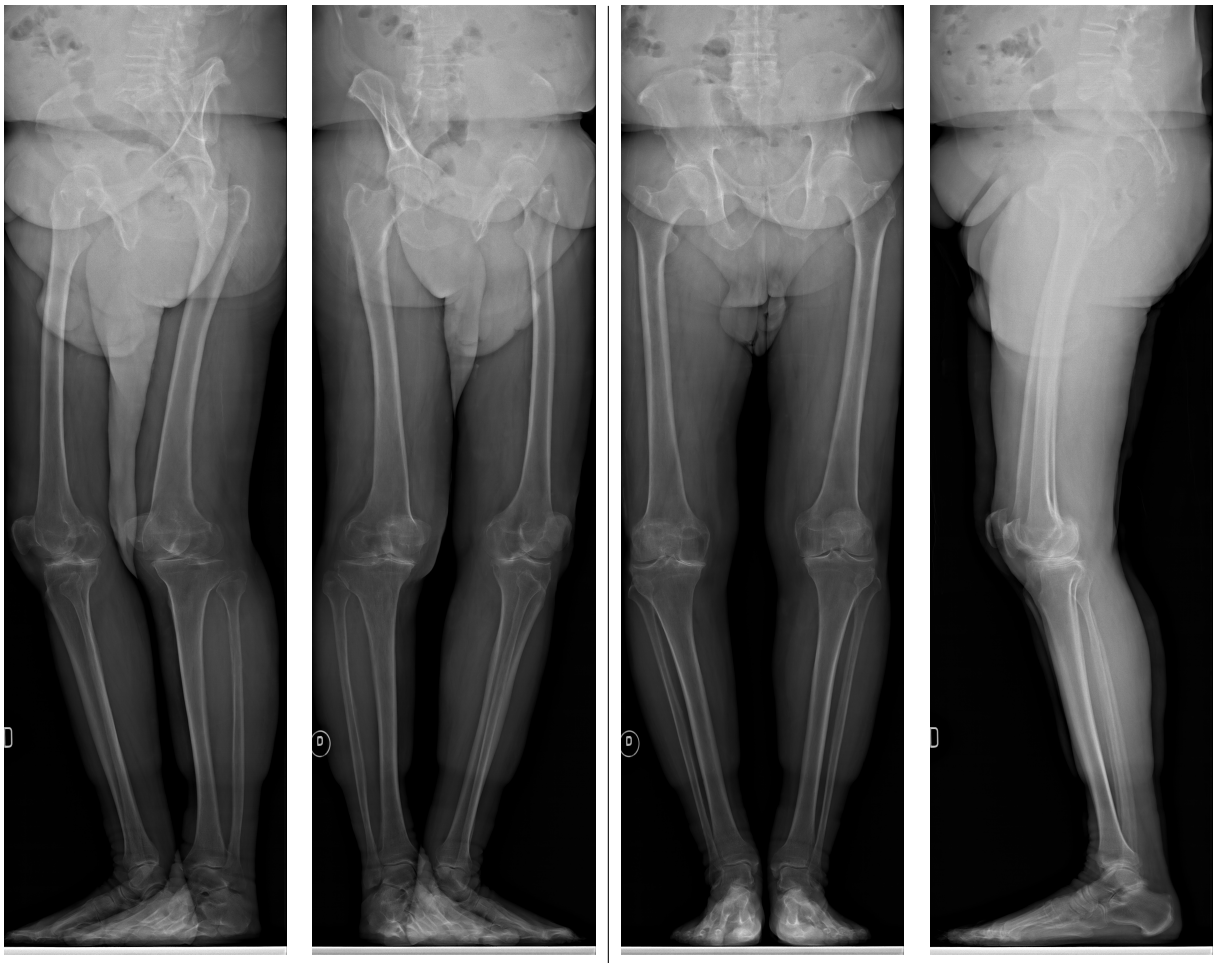


Fig. 29. The pipeline of the proposed method.

Note that most of biplanar 3D reconstruction methods use postero-anterior (PA) and lateral (LAT) images. The human bone structures in the PA image appear distinguished with other bones or dense soft tissue (cartilage, meniscus, and fascia). But in the LAT image, due to the overlap of the bone with the bone itself, other bones or dense soft tissue, very few information about bone structure can be extracted. In contrast, the bone structures in both oblique images appear less clear than in a PA image, but much clearer than in a LAT image. Bone structure information can be equally extracted from two oblique images (see Fig. 30).



45-degree view | 135-degree view | PA view | Lateral view

Fig. 30. Example of biplanar images.

A robust approach will consist of a segmentation method which makes it possible to take into account both the notions of contour and region (detection of homogeneous zones). The cooperation between segmentation by region and contour contributes to a better consideration of the characteristic bones in the radiographic images and, consequently, to a better segmentation [**123**] thanks to the complementary nature of these two types of information

[**177**]. Thereby, segmentation by region-contour cooperation can be expressed as mutual aid between these two concepts in order to improve the end result of segmentation [**42, 120**].

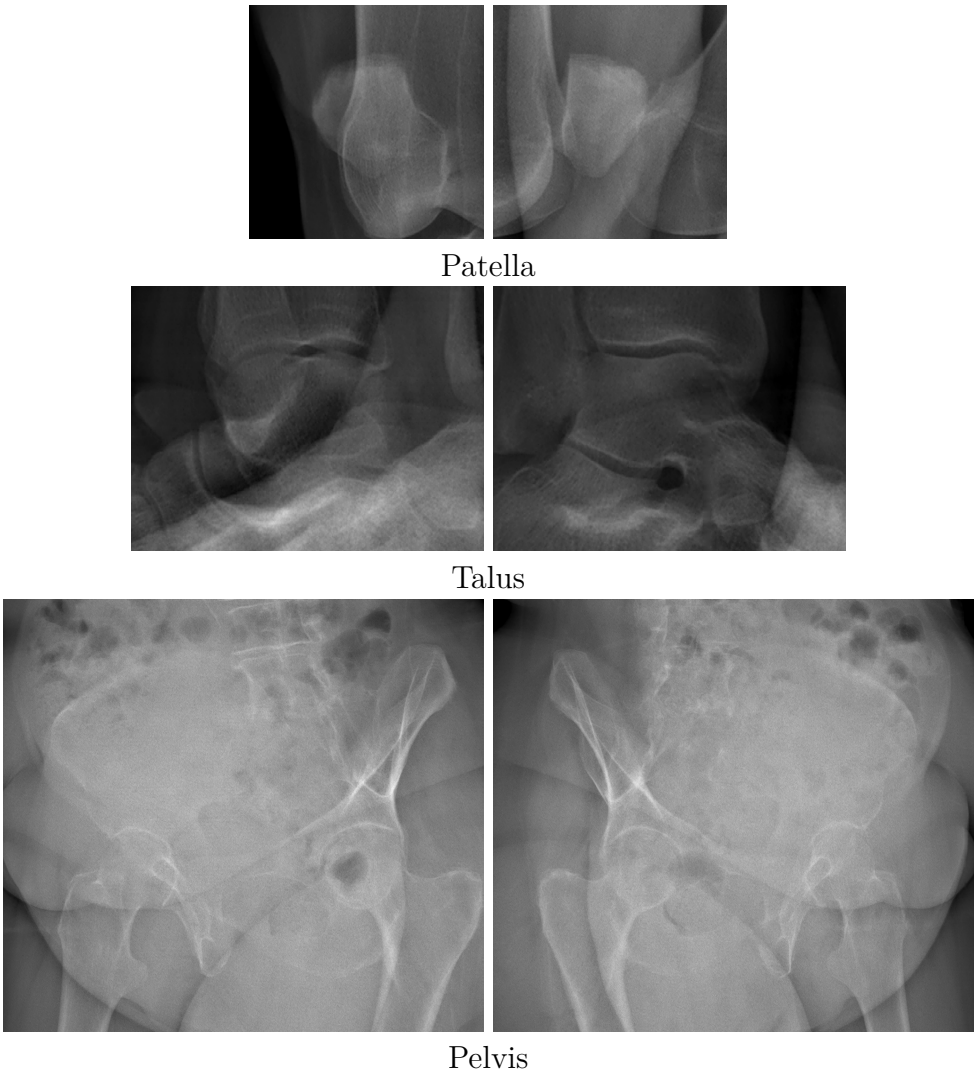


Fig. 31. Example of bone images after initialization (see Section 2.5.1) step. The 45-degree view is on the left and 135-degree view is on the right of each bone.

The prior information is obtained by applying FastMap [**55**], a variant of the MultiDimensional Scaling (MDS) algorithm, on the training set (in which each 3D shape is represented by a fixed-length vector, representing the coordinates of its different points). This algorithm uses an adequate distance metric to find a low enough dimensional space that preserves the chosen metric in the original space. This nonlinear dimensionality reduction step allows us to define a mesh of deformations in a reduced dimensionality space in which each point corresponds to an element of our training set. In our application, this mesh also defines a

concise nonparametric model of the possible 3D deformations in which the admissible statistical deformations, extracted from the training set, are given by the set of values of the different reduced parameter vectors.

From this triangular mesh of nonlinear deformations which summarizes all the *a priori* admissible statistical structures, the objective of the proposed approach is to (non-parametrically) sample a population of possible (*i.e.*, statistically admissible) deformed 3D shapes to be reconstructed. To this end, a sampling strategy is performed from this density mesh by using the interpolation of the nearest (and not dimensionaly reduced) k closest elements of this mesh. An individual fitness value for each candidate shape, sampled from this population, is calculated from the projected contours and Regions of interest (ROIs) on the biplanar oblique radiographic images with the corresponding bone in the preprocessed images (and once after the deformation shape is adjusted by a deterministic optimization algorithm). Finally, the set of fitness values for this population is then used to effectively and iteratively guide a stochastic optimization process to promising solutions until the optimal reconstructed shape is found.

2.1. Training set

We have 3D databases of size 654 for the patella, size 380 for the talus and size 39 for the pelvis. These databases were obtained from the semi-automated segmentation of CT images [**72**]. This semi-automated segmentation method was based on the propagation of the contours in adjacent slices. An initial contour was outlined by manually selecting few high curvature points and using the Fourier interpolation method to complete the contour. Then, a deformable image registration method was applied to map the contour to adjacent slices. Next, for each bone, the surface models were registered by a rigid registration approach called Iterative Closest Point (ICP) [**138**] that iteratively minimizes the distance metric in the least squares error. Finally, we calculated the corresponding points between the models of a bone using descriptors, namely, a set of features, assigned to a point and describing the local geometry around it [**90**]. For each bone, all the models had the same $\mathcal{N}$ number of 3D points. These points were indexed in such a way that the detection of ROIs was easy for all bones.

Fig. 32 represents the regions involved in the segmentation process of the patella, the talus and the pelvis respectively. These ROIs were defined by an orthopedic surgeon.

2.2. Image Preprocessing

Contour detection. The first part of the preprocessing process is performed on each image to enhance the visibility of the component contours which, in fact, constitutes the most important and reliable low-level visual cue in each radiographic image. First, a histogram

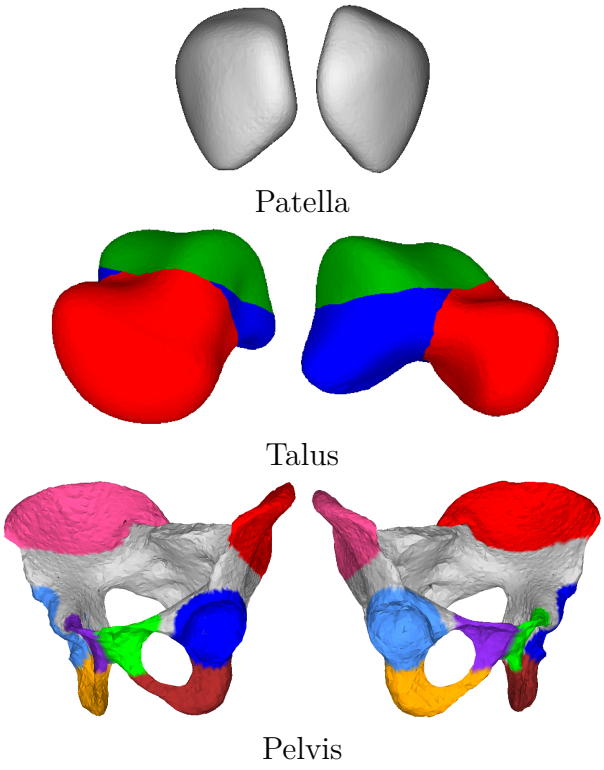


Fig. 32. Example of 3D bone regions. 45-degree view is on the left and 135-degree view is on the right of each bone.

equalization technique increases the global contrast. Second, a median filter and non-local means denoising algorithm [**28**] is used to remove the noise from the images. The non-local means denoising method replaces a pixel with a weighted average of pixels having a similar neighborhood. More precisely, for each pixel, it first searches in a large search window (centered on the pixel to be denoised) for all the neighborhoods of pixels that most closely resemble (with a least squares (LSQ) similarity measure) the neighborhood of the pixel to be filtered. Then, a weighted average (based on the previous LSQ similarity measure) of all these central pixels (of all these neighborhoods) allows to estimate the denoised greyscale value of the pixel. Finally, the edges are detected by using a Canny edge filter [**30**] (see Fig. 33).

Region detection. The next part of the preprocessing step is the region extraction. By using the superpixel and multi-atlas based algorithm [**123**], the input images are segmented into m ROIs. This algorithm registers each image in the training set to the input image using a contour-based registration technique. A superpixel map is estimated from the first half of the optimally registered images in the training set by simply taking the intersection of all the regions existing in this selected subset. Each superpixel represents the (local) variability of bone regions. This map is adaptive to the input image and takes into account all the nonlinear and local variability of bone regions existing in the selected subset. Then a pruning step

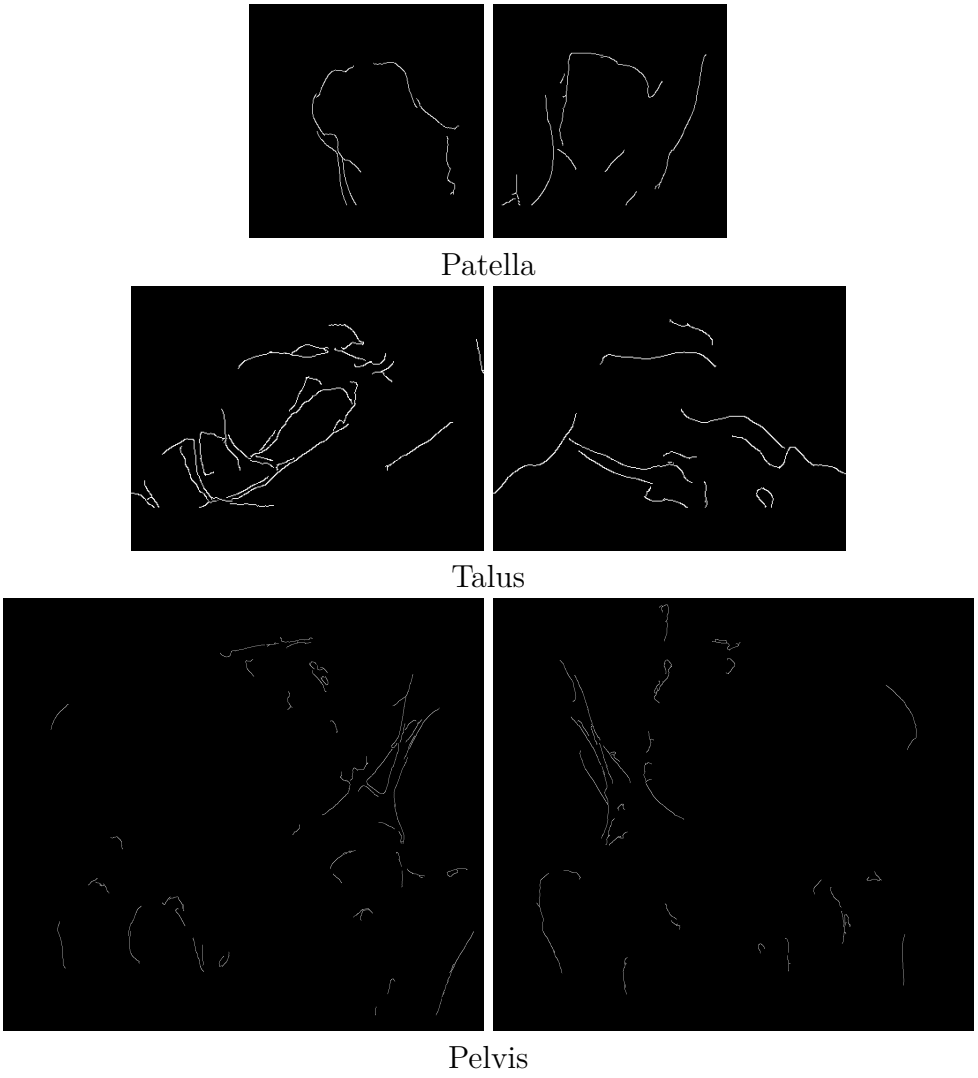


Fig. 33. Example of bone contour detection. 45-degree view is on the left and 135-degree view is on the right of each bone.

is achieved by finding the set of connected superpixels which maximize the contour-based similarity between the outer contour of this superpixel map and the edge map of the input image. Finally, a label propagation step based on the entropy concept is used for refining the resulting segmentation map into the most likely internal regions to the final consensus segmentation (see Fig. 34).

2.3. Statistical Deformable Model

This MDS based algorithm has been successfully used in many imaging applications, such as 3D object recognition [**36**], human action recognition [**158, 16**], image segmentation [**107, 108**], and image change detection [**159**]. In our study, each object in the training set is considered to be a point in $\mathcal{N}$-dimensional space. These points are then projected on $\mathcal{M}$

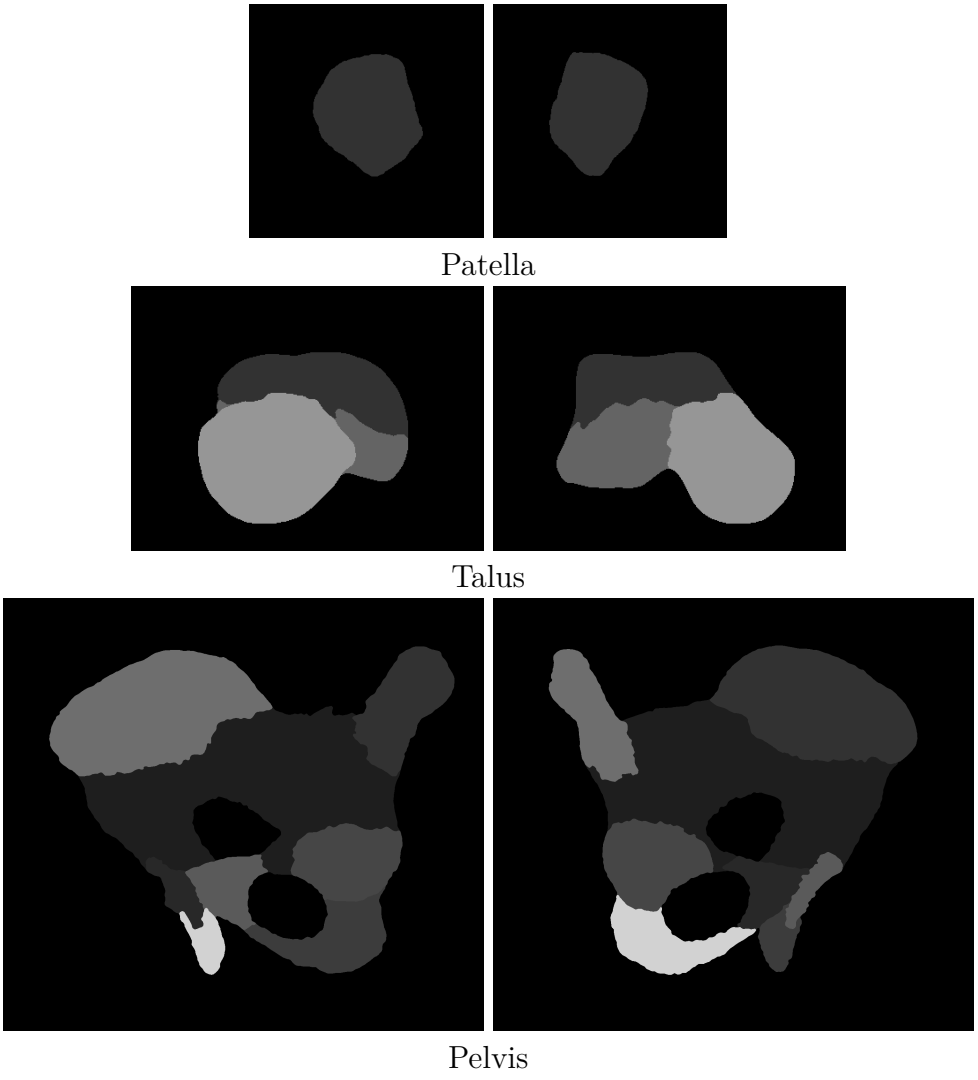


Fig. 34. Example of bone region detection. The 45-degree view is on the left and 135-degree view is on the right of each bone.

mutually orthogonal axes. The goal is to calculate the training set in an $\mathcal{M}$-dimensional space from the distance matrix previously computed in the original space ($\mathcal{M} << \mathcal{N}$). FastMap algorithm. is based on the projections of the objects on a selected line. This line is created by connecting two furthest pivot objects O_a and O_b in the training set. Then the projections of the objects O_i are computed by applying the cosine law defined as:

$$\mathcal{D}^2(O_b,O_i) = \mathcal{D}^2(O_a,O_i) + \mathcal{D}^2(O_a,O_b) - 2x_i\mathcal{D}(O_a,O_b) \tag{2.1}$$

where $\mathcal{D}(O_i,O_j)$ (for $i,j = 1,...,\mathcal{N}$) is the distance between O_i et O_j. From the Pythagorean theorem, Eq. 2.1 can be solved for x_i, the first coordinate of object O_i:

$$x_i = \frac{\mathcal{D}^2(O_a,O_i) + \mathcal{D}^2(O_a,O_b) - \mathcal{D}^2(O_b,O_i)}{2\mathcal{D}(O_a,O_b)} \tag{2.2}$$

For $\mathcal{M} > 1$, the objects are projected on a hyper-plane $\mathcal{H}$ which is perpendicular to the line O_aO_b. A new distance $\mathcal{D}'$ between two projections O'_j et O'_j on $\mathcal{H}$ is deduced from the original distance $\mathcal{D}$ as follows:

$$\mathcal{D}'^2(O'_iO'_j) = \mathcal{D}^2(O_iO_j) - (x_i - x_j)^2 \tag{2.3}$$

Then Eq. 2.2 (with new distance function $\mathcal{D}'$, and new pivot objects) is applied to obtain the next coordinate of objects (i.e. dimension $\mathcal{M} > 1$).

Algorithm 6 Heuristic to choosing two distant objects

begin

- Chose arbitrarily an object, and declare it to be the second pivot object O_b
- $O_a \leftarrow$ the farthest object from O_b according to the distance function $\mathcal{D}$
- $O_b \leftarrow$ the farthest object from O_a according to the distance function $\mathcal{D}$
- return O_a and O_b as the desired pair of object

end

The efficiency of FastMap, in terms of preserving information in reduced dimensionality, can be evaluated by calculating a correlation metric [**74**]. This metric is the correlation of Euclidean distance of each object pair in the original space and their corresponding pair in reduced dimensionality space:

$$\text{correlation}(X,Y) = \frac{\text{covariance}(X,Y)}{\sigma_X\sigma_Y} = \frac{\frac{X^tY}{|X|} - \overline{XY}}{\sigma_X\sigma_Y} \tag{2.4}$$

where X, Y are respectively the vector of distance in the original space and in the reduced dimensionality space. X^t, $|X|$, $\overline{X}$ and σ_X are the transpose, the cardinal, the mean, and the standard deviation of X, respectively.

This correlation specifically quantifies the degree of dependence between X and Y and show how FastMap is able to give a cartography in the reduced dimensionality space in which each point is placed such as the inter-point distances (in the original space) are preserved as well as possible [**40**]. A perfect correlation of 1 gives a perfect (positive) linear correlation or relation between the high dimensional data and the low dimensional data (i.e. no loss of information) and a correlation of 0 means a total loss of information (e.g. a correlation of 0.80 means that the FastMap technique succeeds in keeping 80% identical in term of distance of the object pairs between the two [original and reduced] spaces).

Fig. 35, 36, and 37 show the cartographies or density mappings of patella, talus and pelvis shapes at dimension $\mathcal{M} = 1$ and at $\mathcal{M} = 2$, respectively, and showing for a complex shape like the pelvis, several modes of deformation.

Table 9 shows the correlation metric of patella, talus, and pelvis at different reduced dimensions.

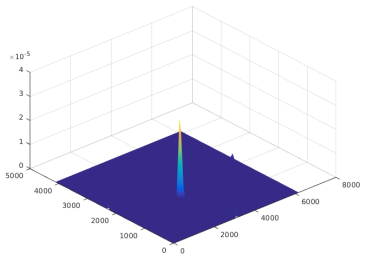
For each possible (or statistically admissible) candidate 3D deformed shape, which will be sampled from the preliminary estimated shape mappings (in reduced dimensionality), the

Algorithm 7 FastMap

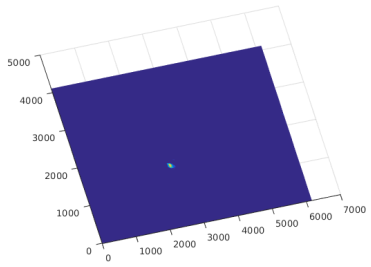
$\mathcal{O}$	set of objects in original dimensional space
$\mathcal{M}$	reduced dimensionality
$\mathcal{Y}$	set of objects in reduced dimensional space

1. Initialization
$t \leftarrow 0$

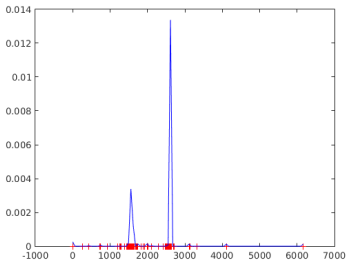
2. Algorithm Fastmap($\mathcal{M}$, $\mathcal{D}()$, $\mathcal{O}$)
if $\mathcal{M} \leq 0$ **then**
  return
end if
$t \leftarrow t + 1$
Choose pivots 0_a, 0_b by Algorithm 6
if $\mathcal{D}(0_a,0_b) == 0$ **then**
  $\mathcal{Y}[i,t] = 0$ for every i
  return
end if
for each $O_i \in \mathcal{O}$ **do**
  Compute x_i using Eq. 2.2
  $\mathcal{Y}[i,t] = x_i$
end for
Call Fastmap($\mathcal{M} - 1$, $\mathcal{D}'()$, $\mathcal{O}$)



(a) $\mathcal{M} = 2$, profile view

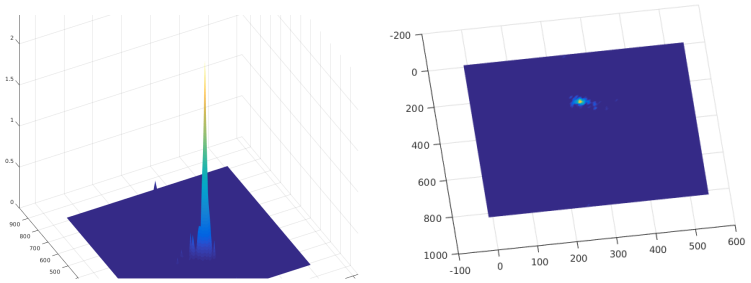


(b) $\mathcal{M} = 2$, top view

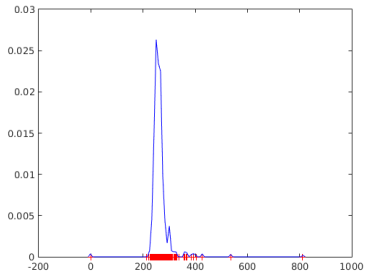


(c) $\mathcal{M} = 1$

Fig. 35. Distribution of patella in low dimensionality space

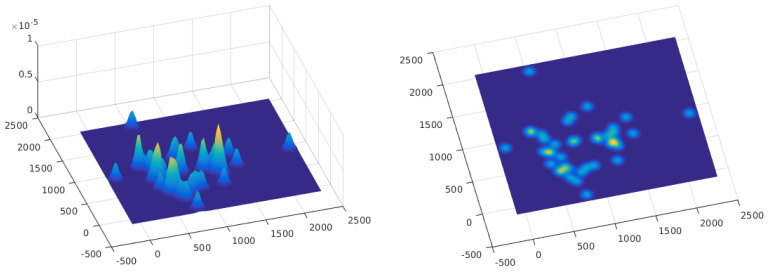


(a) $\mathcal{M} = 2$, profile view (b) $\mathcal{M} = 2$, top view

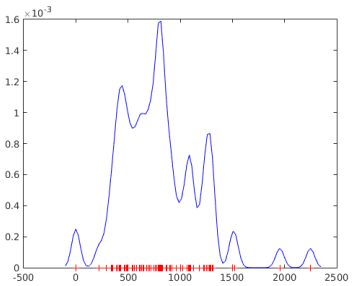


(c) $\mathcal{M} = 1$

Fig. 36. Distribution of talus in low dimensionality space



(a) $\mathcal{M} = 2$, profile view (b) $\mathcal{M} = 2$, top view



(c) $\mathcal{M} = 1$

Fig. 37. Distribution of pelvis in low dimensionality space

	$\mathcal{M}=1$	$\mathcal{M}=2$	$\mathcal{M}=3$
Patella	0.753	0.84	0.896
Talus	0.847	0.92	0.94
Pelvis	0.682	0.734	0.819

Tableau 9. Correlations of patella, talus and pelvis at different reduced dimensions.

proposed model will find the k nearest neighbors of this sample in the low-dimensional data, and then will use these k corresponding objects in the high-dimensional space to generate, by interpolation (by applying the inverse distance weighting (IDW) function), a new instance of the S bone shape. The weight w is an inverse distance of a point to its nearest neighbors and is computed in the reduced dimensionality space. The new bone shape is calculated in the original space.

$$S(z) = \frac{\sum_i^k w_i * Z_i}{\sum_i^k w_i} \tag{2.5}$$

$$w_i = \frac{1}{dist(z,z_i)^{\varrho}} \tag{2.6}$$

where z is a point in the reduced dimensionality space, $dist$ is the distance between 2 points, $\varrho \geq 0$ is the power parameter, and Z is an object in the (original) high dimensional space.

2.4. Energy Function Term

2.4.1. Likelihood Energy Term.

Edge potential field-based similarity.

An edge potential field-based similarity measure evaluates the concordance or the similarity between the external contours of the bone silhouette projections on the two biplanar X-ray images and an edge potential field, calculated from the previously detected contours. This edge potential field attracts the bone contours and aligns them on the edge of the input image, by giving (concretely) a similarity measure all the greater as the projected contours coincide well with the edges existing in the images. In addition, a directional component is added to complete the measurement of the correspondence between the projected contours of the human bone and the edges in the two views. This measure is computed on the preprocessed image and is defined as [**75**]:

$$\xi(S) = \frac{1}{n}\sum_n(\Phi(x,y)|\cos(\alpha(x,y))|) \tag{2.7}$$

$$\Phi(x,y) = \exp(-\rho\sqrt{\delta_x^2+\delta_y^2}) \tag{2.8}$$

where (δ_x,δ_y) is the displacement to the nearest edge point in the image, ρ is a smoothing factor which controls the degree of smoothness of the potential field Φ, $\alpha(x,y)$ is the angle between the tangent direction of the projected external contours at (x,y) and the tangent of

the nearest edge, and n is the number of pixels on projected external contours of the human bone S (see Fig. 38).

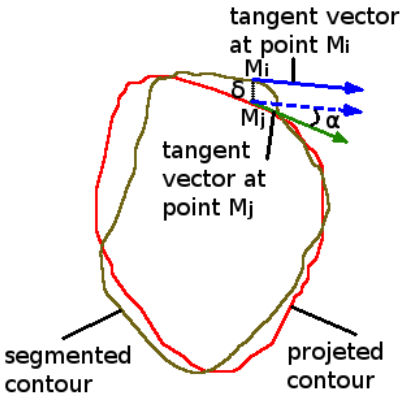


Fig. 38. Directional component used in equation (2.7).

Non-overlapping ratio-based similarity.

The non-overlapping ratio evaluates the matching between regions of the segmented image and the bone silhouette projection on biplanar X-ray images. This ratio is simply the number of pixels of segmented and projected regions that are missing from their intersection on the number of pixels of their union. A ratio of 0 means a perfect overlap of two regions and 1 means completely dissimilarity. This ratio for m ROIs is defined as [**76**]:

$$\vartheta(S) = \prod_{i=1}^{m} \left(\frac{|R_i \cup R_i(S)| - |R_i \cap R_i(S)|}{|R_i \cup R_i(S)|} \right)^{\frac{1}{m}} \tag{2.9}$$

where S is the 3D bone model, R_i is the set of pixels in i^{th} ROI in segmented image, $R_i(S)$ is the set of pixels of the projection of i^{th} ROI in 3D model S and $|R|$ indicates the number of pixels in region R (see Fig. 39).

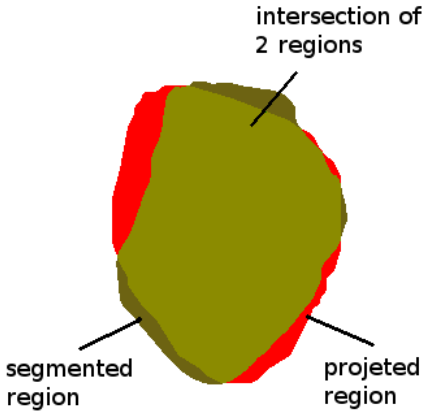


Fig. 39. Overlap of segmented and projected regions

The likelihood energy term is defined as:

$$E_l(S) = (1 - \xi(S)) + \vartheta(S) \tag{2.10}$$

2.4.2. Prior Energy Term. In order to favour statistically admissible surface structures that often appear in our training base, our prior energy term will favour a candidate shape in a high density area of our shape density mapping and will be simply defined as the average distance between it and its k nearest neighbors:

$$E_p(s) = \frac{1}{k} \sum_{i}^{k} (s - s_i)^2 \tag{2.11}$$

where s_i is a point in the k nearest neighbourhood of s in the reduced dimensionality space.

2.5. 3D Reconstruction

2.5.1. Initialization. From biplanar oblique (45 and 135 degree) X-ray images (I_{45} and I_{135}), we construct a box which contains the bones to be reconstructed. The dimensions of the box are the width of the 45-degree image, the width of the 135-degree image and the height of the highest image. Then, we manually position a 3D model of the bones (patella, talus, pelvis) in this box in such a way that the projections of these 3D models coincide with their corresponding bones in the two images. This ref_{box} box constitutes a reference box and these bone models ($ref_{patella}, ref_{talus}, ref_{pelvis}$) constitute reference models. These references are computed offline and only once.

The initialization step of our approach is simply the registration of ref_{box} on in_{box} which is created by using the aforementioned method on biplanar input images. The transformation from this registration is then applied to the ref_{box} and the reference models ($ref_{patella}, ref_{talus}, ref_{pelvis}$). Fig. 40 well illustrates the initialization step and Fig. 31 shows a result example. This step can be refined by using a 3D/2D registration. The registered reference models will be used for the next step.

2.5.2. Optimization. Our method is based on the optimization of an objective function E that contains both a contour/region based similarity term E_l and a shape prior term E_p:

$$E(S(\theta)) = E_l(S,I_{45}) + E_l(S,I_{135}) + \beta E_p(s) \tag{2.12}$$

where θ is the deformation parameters, S and s are the bone shape in the original and reduced dimensionality space, β is a weighting factor which indicates the importance of the prior information.

To minimize this complex non-convex energy function, we resort to Exploration Selection (ES) algorithm, a stochastic and efficient optimization algorithm, proposed in [**60**]. This algorithm belongs to the class of evolutionary algorithms, and is typically well-suited for this type of function to be optimized. This class of algorithm is inspired by the evolution of

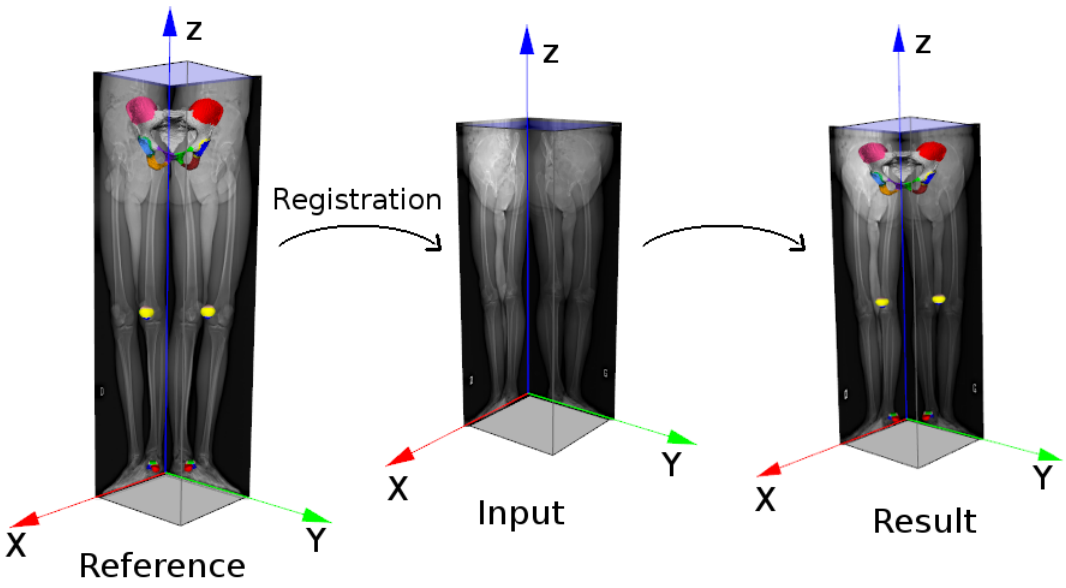


Fig. 40. Initialization step

nature by reflecting the process of natural selection to provide good approximate solutions to complex problems. These algorithms have been successfully used in medical imaging [**18, 127, 115**].

The ES algorithm can be summarized in two steps: exploration and selection steps. Let F be a finite discrete subset of the Cartesian product of h intervals $[m_i,M_i]$ for $1 \leq i \leq h$, and $\theta = \{\theta_1,...,\theta_n\}$ a set of n candidate solutions (population) randomly chosen. $\hat{\theta}$ is the optimal element θ_i of θ such that $E(\theta_j) > E(\theta_i)$ for $1 \leq j < i$, and $E(\theta_j) \geq E(\theta_i)$ for $i < j \leq n$. In the first step, each solution of θ is considered as an individual that attempts a random search on the exploration graph. The exploration process acts independently on each individual, and chooses a random f according to a positive distribution. We compute $\hat{\theta} = \arg\min_\theta E(.)$, and for $i \leq f$, we replace θ_i by $\gamma_i \in N(\theta_i) \setminus \{\hat{\theta}\}$ according to a uniform distribution, where $N(b)$ is defined by $\{a \in F :$ for some $j, |a_j - b_j| \leq r(M_j - m_j), a_i = b_i, i \neq j\}$, and $r \in [0,1]$ is the radius of exploration. Otherwise, θ_i is changed by $\hat{\theta}$ in the second step. This process is run until a stopping criterion has been met (cf. Algorithm 3).

2.5.3. Refinement. We use an additional strategy which consists in associating the stochastic exploration, selection search with a local optimization technique. In each generation (i.e., each Exploration/Selection step), the best individual or solution is used to initialize a gradient descent-style deterministic minimization technique in which local deformations are used to refine the bone shape result previously estimated by the stochastic optimizer (at each E/S step). Therefore, the best individual deterministically explores the local neighborhood in parameter space to find a point that further minimizes our energy function. More precisely, this deterministic gradient descent procedure uses local deformations to refine the

Algorithm 8 ES optimization algorithm

$E(.)$	A h-variable function defined on F to be minimized
$N(a)$	The neighborhood of an element $a \in F$ defined by $\{b \in F$: for some $1 \leq j \leq l$, $\lvert b_j - a_j \rvert \leq r\,(M_j - m_j)$, $b_i = a_i$, $i \neq j\}$
D	$(= h/r)$ The diameter of the exploration graph F (endowed with the neighborhood system $\{N(a)\}_{a \in F}$)
θ	$\theta = (\theta_1, \ldots, \theta_n)$, an element of F^n (population)
$\hat{\theta}$	$\hat{\theta} \in F$, $\hat{\theta} = \arg\min_{\theta_i \in (\theta_1, \ldots, \theta_n)} E(\theta_i)$ *i.e.*, the minimal point in θ with the lowest label
p	The probability of exploration
t	The iteration step
F	A finite discrete subset of the Cartesian product $\Pi_{j=1}^{h}[m_j, M_j]$ of h compact intervals
n	The size of the population (greater than D)
r	A real number $\in [0,1]$ called the radius of exploration (with r greater than the ϵ-machine)

1. Initialization
Random initialization of $\theta = (\theta_1, \ldots, \theta_n) \in F^n$
$t \leftarrow 2$

2. Exploration/Selection
while a stopping criterion is not met **do**
(1) Compute $\hat{\theta}$; $\hat{\theta} \leftarrow \arg\min_{\theta_i \in \theta} E(\theta_i)$
(2) Draw f according to the binomial law $b(n,p)$
- For $i \leq f$, replace θ_i by $\gamma_i \in N(\theta_i) \setminus \{\hat{\theta}\}$ according to the uniform distribution (*Exploration* step)
- For $i > f$, replace θ_i by $\hat{\theta}$ (*Selection* step)

(3) $t \leftarrow t + 1$ and $p \leftarrow t^{-1/D}$
end while

bone shape result obtained with θ previously estimated. This procedure can be described as follows:

- Do
  - For each point p_i on external contour of bone shape $S(\theta)$
    - Find K_i nearest neighbour of p_i
    - Compute the normal η_i which is the average of the normal of K_i
    - Compute $\vartheta(S(\theta, K_i))$, $\vartheta(S(\theta, K_i - \varepsilon\eta))$, $\vartheta(S(\theta, K_i + \varepsilon\eta))$
    - Update K_i with the lowest energy ϑ
- While the energy ϑ is not stable.

3. Experiments

3.1. Lower Limb (Patella, Talus, Pelvis) Database

Our training databases consist of 654 surface models for the patella, 380 for the talus and 39 for the pelvis. These training surface models were constructed from binary volumes that were semi-automatically segmented from CT-datasets of patient [**72**].

3.2. Radiographic images

In our application, we used biplanar oblique (45 and 135 degree) X-ray images (I_{45} and I_{135}) acquired with a low dose imaging device. This system performs the simultaneous acquisition of two orthogonal calibrated X-rays, with the patient in a standing position. The radiographic image illustrates the superposition of the various structures on the same plan and gives place to semi-transparent images. The size of our radiographic images is around 1764×5932 pixels (coded on 256 gray levels).

3.3. Calibration

Calibration is a necessary step to compute the geometrical parameters of the radiological environment. The images are calibrated by using images of a planar grid acquired in different orientations. First, a homography which represents the mapping between the grid points and image points was estimated. Then the parameters of the projection matrix were extracted form the homography. The projection of the 3D point (X, Y, Z) is defined as:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & x_0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ YZ \\ Z \end{pmatrix} \tag{3.1}$$

where $(x, y, 1)$ is the projection point, f is the focal length, x_0 is the optical center and s is scaling factor [**50**].

3.4. Comparison protocol

Our 3D reconstruction method was validated on 13 bones (5 patellae, 5 tali, and 3 pelvis) from 13 patients (13 pairs of radiographic images ((I_{45} and I_{135} views) of the lower limbs). This comparison was made using the distance (root mean square (RMS), and maximum) between the point from the reconstructed lower limb and the nearest point of the corresponding lower limb obtained with CT-Scan, which was considered as the ground truth and whose accuracy is ± 1 mm for the human spine [**10**].

3.5. Experimental results

In our study, we have used the shape density mapping, obtained by the FastMap, in the reduced space of dimension 3, that preserves 94% information for talus and more than 80% information for pelvis and patella in terms of distance of object pairs.

Based on preliminary test results, we chose k = 12, 8, 18 for the nearest neighbour and ϱ = 1.2, 2.8, 0.8 for the power parameter in the IDW interpolation for the patella, talus, and pelvis respectively. The prior weighting factor β was set to 0.001, 0.0005, 0.00005 for the patella, talus, and pelvis respectively to weight the prior energy term with respect to the likelihood energy term. Finally, we set the size of the population to 100 and the number of iterations to 380 for ES algorithm.

We have validated our method on 5 patellae, 5 tali and 3 pelvis by using the aforementioned protocol. The results of comparisons are given in Table 10, 11, and 12. Fig. 41 details the errors of each model of each bone. Fig. 42, 43, and 44 show examples of the reconstructed bone by our method and theirs projections on both images.

Model	**RMS** (mm)	**Maximum error** (mm)	**Errors < 1mm** (%)	**1mm ≤ Errors < 2mm** (%)	**Errors ≥ 2mm** (%)
1	0.8	2.6	81.4	18.2	0.4
2	0.9	3.3	74.9	21.9	3.2
3	0.6	2.0	89.7	10.3	0.0
4	1.0	3.0	69.6	26.7	3.7
5	1.3	4.1	59.9	28.4	11.7
Average ± SD	0.9 ± 0.2	3.0 ± 0.7	75.1 ± 10.1	21.1 ± 6.5	3.8 ± 4.2

Tableau 10. Results on comparison of 5 patellae.

Model	**RMS** (mm)	**Maximum error** (mm)	**Errors < 1mm** (%)	**1mm ≤ Errors < 2mm** (%)	**Errors ≥ 2mm** (%)
1	1.7	6.5	48.1	29.8	22.1
2	1.2	3.9	56.4	36.4	7.2
3	1.3	4.4	57.8	31.2	11.0
4	1.5	4.8	44.1	39.8	16.1
5	1.5	3.9	51.5	31.3	17.2
Average ± SD	1.4 ± 0.2	4.7 ± 1.0	51.6 ± 5.1	33.7 ± 3.8	14.7 ± 5.2

Tableau 11. Results on comparison of 5 tali.

Model	**RMS** (mm)	**Maximum error** (mm)	**Errors < 1mm** (%)	**1mm ≤ Errors < 2mm** (%)	**2mm ≤ Errors < 4mm** (%)	**Errors ≥ 4mm**
1	3.6	14.3	25.8	24.9	25.0	24.3
2	2.5	13.9	40.0	25.6	22.8	11.6
3	4.2	19.3	20.8	21.0	29.3	28.9
Average ± SD	3.4 ± 0.7	15.8 ± 2.5	28.9 ± 8.1	23.8 ± 2.0	25.7 ± 2.7	21.6 ± 7.3

Tableau 12. Results on comparison of 3 pelvis.

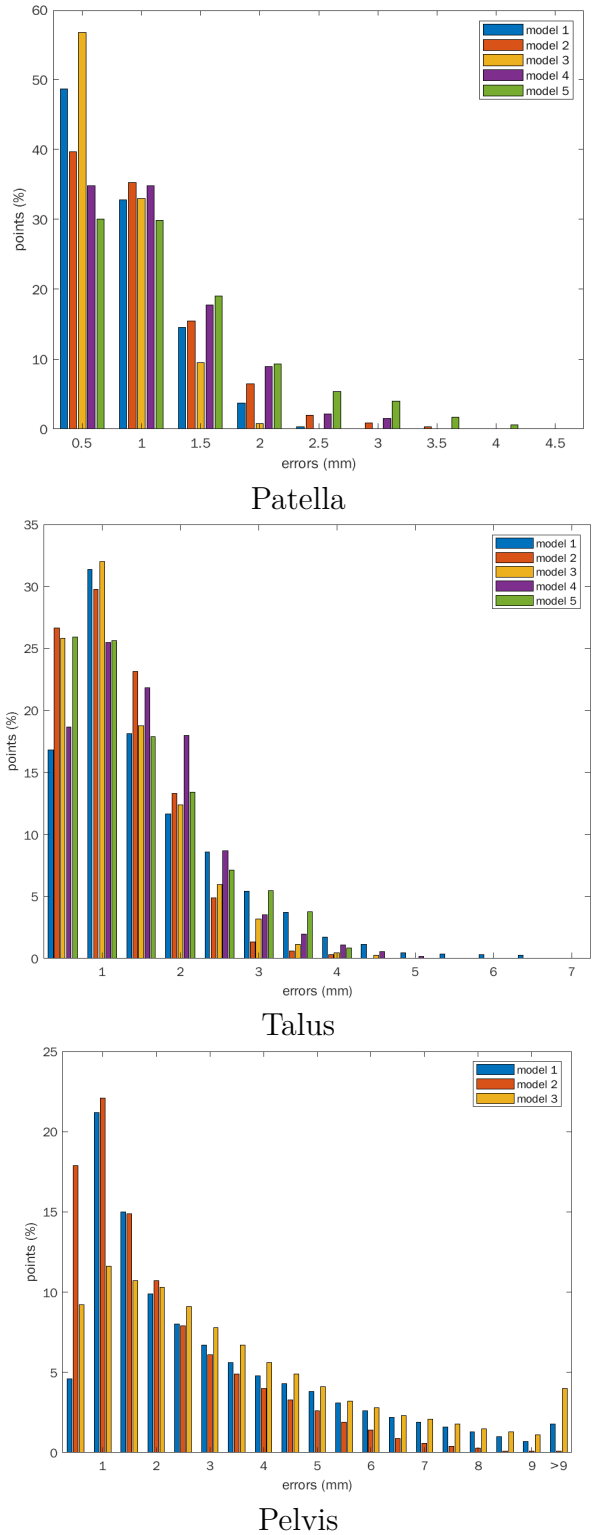


Fig. 41. Detailed errors of each model of each bone.

For the patella, the average RMS was 0.9 mm and the average maximum error was 3.0 mm. 75.1% of points had an error less than 1 mm. 21.1% of points had an error between 1 mm and 2 mm. The error of the remaining 3.8% of points was greater than 2 mm.

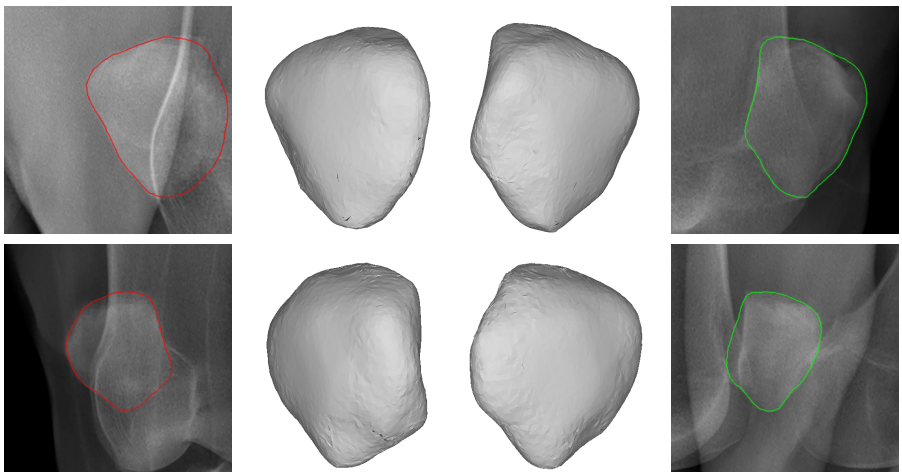


Fig. 42. Examples of the reconstructed patellae by our method and their projections on both images.

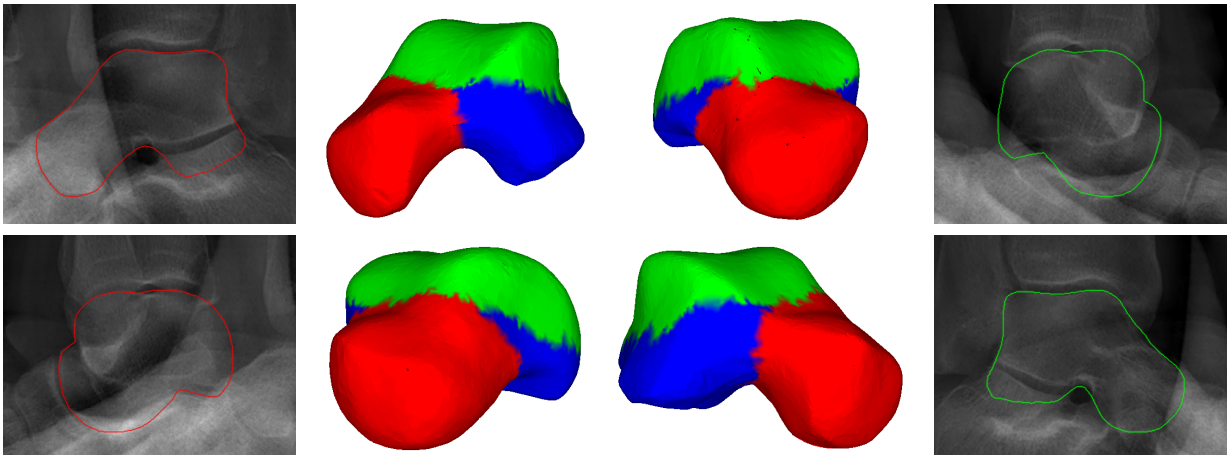


Fig. 43. Examples of the reconstructed tali by our method and their projections on both images.

For the talus, the average RMS was 1.4 mm and the average maximum error was 4.7 mm. 51.6% of points had an error less than 1 mm. 33.7% of points had an error between 1 mm and 2 mm. The error of the remaining 14.7% of points was greater than 2 mm.

For the pelvis, the average RMS was 3.4 mm and the average maximum error was 15.8 mm. 52.7% of points had an error less than 2 mm. 25.7% of points had an error between 2 mm and 4 mm. The error of the remaining 21.6% of points was greater than 4 mm.

4. Discussion and Conclusion

In this paper, we have presented a new approach for the 3D reconstruction of human lower limbs (patella, talus, and pelvis) from two radiographic oblique projections (45 and 135 degrees). This approach efficiently exploits all the information existing in the image (edges

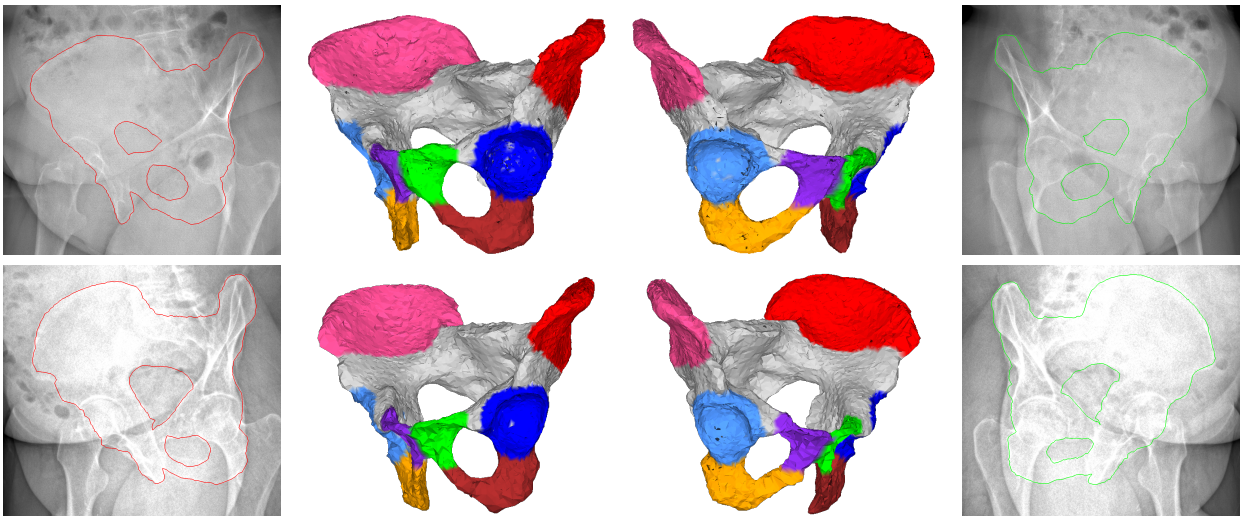


Fig. 44. Examples of the reconstructed pelvis by our method and their projections on both images.

and ROIs) and the estimation of a FastMap both for the reduction of the dimensionality and to constrain the 3D reconstruction problem. The proposed 3D reconstruction problem is considered as a cost function optimization problem encompassing the information extracted from the image and the information extracted from the database by FastMap. Our 3D reconstruction method has the advantage of minimizing the X-ray dose (only two oblique radiographic images), exploitation of the maximum amount of information existing in these two images (contours and ROIs) and the use of the nonlinear FastMap statistical method which gives a good representation of deformations in the bone database relative to those used in other 3D reconstruction methods. The results obtained in our experiments are quite encouraging and indicate that accurate unsupervised 3D reconstruction is technically feasible.

This approach has been validated on a sample of 13 bones (5 patellae, 5 tali, and 3 pelvis) from 13 patients (13 pairs of radiographic images (I_{45} and I_{135} views) of lower limbs), by comparing the model obtained from our approach and those obtained with CT-Scan, which was considered as the ground truth and whose accuracy is ± 1 mm for the human spine [**10**]. The average RMS was 0.9 mm, 1.4 mm, and 3.4 mm for the patella, the talus, and the pelvis respectively. The mean and the standard deviation of the percentages for the points whose error is less than 1 mm for the patella and the talus were $75.1 \pm 10.1\%$ and $51.6 \pm 5.1\%$. Those whose error is less than 2 mm for the pelvis were $52.7 \pm 9.8\%$.

The experiments showed that our results are better than that of the hierarchical statistical modeling method [**18**] in case of the patella and talus (see Table 13). The errors of the pelvis reconstruction were greater than other bone reconstructions because the pelvis has a very complex structure and its database was smaller than that of the other bones, limited to 39

Our method			Hierarchical statistical modeling method		
Bones	**Average RMS** (mm)	**Average maximum error** (mm)	**Bones**	**Average RMS** (mm)	**Average maximum error** (mm)
Patella	0.9	3.0	**Thoracic vertebra**	1.6	4.5
Talus	1.4	4.7	**Lumbar vertebra**	1.9	5.4
Pelvis	3.4	15.8			

Tableau 13. Results of our method versus the hierarchical statistical modeling method [**18**].

objects in our application. Also, the sacrum and the coccyx were blurred in the images in addition to the fact that it's difficult to distinguish the sacrum from the spine. Note that the structural complexity, the anatomical position in the human body, and the number of elements in the training set of each bone of each method are different.

Compared to the biplanar method proposed in [**18**] which uses only the contours of anatomical shapes, our likelihood function uses all the information of the two images (contours and ROIs). Moreover, the linear PCA-based dimensionality reduction used in [**18**] doesn't perfectly represent admissible statistical deformations existing in the training set because linearity is a hypothesis that is not necessarily true in our context.

Recall also that the biplanar technique presented in [**18**] is also limited due to the inherent inaccuracy produced in the segmentation of I_{LAT} (leading to reconstruction errors). In addition, this method does not use all the information contained in two X-ray projections, for example, the contours of each bone structure and the geometric structure or statistical knowledge of the possible deformation of the bone structure to be reconstructed.

The proposed approach requires a training representative database. Nevertheless, the proposed reconstruction method remains unsupervised in the sense that this database is constructed *off-line* and not during the 3D reconstruction step.

Our approach provides an accurate representation of the patella, talus, and pelvis from just two x-ray views, while the CT-Scan requires hundreds of images to achieve the same three-dimensional reconstruction with a similar precision level. The proposed method is therefore interesting in terms of the quantity of data to be acquired, processed, and managed.

Our proposed scheme thus constitutes an alternative to CT-Scan 3D reconstruction with the advantage of low irradiation and will be of great interest for diagnosis of bone structure deformities, simulation of orthopedic treatments, and for reliable geometric models for finite element studies. However, in the moment, this reconstruction method is not suitable without improvement for surgical navigation applications when compared to CT-Scan reconstruction errors of ± 1 mm, the current standard for those applications.

To our knowledge, the proposed approach was the first 3D reconstruction method for the talus and remains sufficiently general to be applied to other medical reconstruction problems for which a database of the anatomical structure is available (with two radiographic views).

In addition, this method, based on the ES optimizer, is easily parallelizable and thus remains especially well-suited for the next-generation GPU or massively parallel computers and multi-core processors.

Discussion générale

Nous avons présenté dans ce travail de recherche appliquée une méthode de segmentation des os, une méthode de recalage 3D/2D des composants de la prothèse du genou, et une méthode de reconstruction 3D des os. Toutes ces méthodes fonctionnent avec les images radiographiques biplanaires obliques. Notez que la plupart des méthodes biplanaires utilisent des images radiograpiques PA et LAT. Les structures osseuses humaines dans l'image PA se distinguent adéquatement des autres os ou des tissus mous denses (cartilage, ménisque et fascia). Mais dans l'image LAT, en raison du chevauchement de l'os avec lui-même, d'autres os ou des tissus, très peu d'informations sur la structure osseuse peuvent être extraites. Dans les deux images obliques, les structures osseuses apparaissent moins claires que dans l'image PA mais beaucoup plus claires que dans l'image LAT. Les informations sur la structure osseuse donc peuvent être également extraites de deux images obliques.

La méthode de segmentation des os a été validée sur un ensemble de données de 31 images radiographiques pour chaque structure osseuse avec une précision de classification, en comparant avec les segmentations manuelles (étalon-or), de 93,79% pour la rotule, de 88,30% pour l'astragale et de 85,02% pour le bassin. Nous avons aussi comparé notre méthode avec les méthodes PB et MV en utilisant différentes stratégies de recalage (recalage affine et notre recalage basé sur la forme). Notre méthode est légèrement meilleure que les autres méthodes (MV et PB) avec deux stratégies de recalage différentes dans le cas d'os simples (rotule, astragale), et beaucoup meilleure dans le cas d'une structure complexe (bassin). Notez que notre méthode est proposée pour l'imagerie radiographique à rayons X et les autres méthodes sont utilisées pour la MRI. À cause des différents bruits quantiques dans différents systèmes d'imagerie radiographique, de nombreux artefacts existent dans ces images, et principalement en raison du fait que les tissus voisins à l'intérieur du corps humain peuvent avoir des taux d'absorption des rayons X similaires, ces phénomènes peuvent considérablement biaiser les estimations des autres méthodes. Le point fort de notre approche réside dans la combinaison de plusieurs informations faibles mais complémentaires, tels que des mesures de similarité basées sur les contours et sur les régions tout au long des différentes étapes de notre processus. En plus, notre approche basée sur des superpixels, qui prend en compte toute la variabilité non linéaire et locale des régions osseuses existant dans la base d'apprentissage ainsi que

la combinaison de l'étape de fusion de propagation d'étiquettes, au sens de la variation de l'information, et l'étape de suppression des valeurs aberrantes, nous permet de contraindre efficacement l'espace des candidats de segmentation, pour produire finalement une carte de segmentation fiable même à partir des images radiographiques de faible qualité.

La méthode de recalage 3D/2D des composants de la prothèse du genou a été validée sur un ensemble de 64 composants (32 composants fémoraux et 32 composants tibiaux) et 64 couples d'images radiographiques synthétiques avec différentes mesures de similarité. Les RMSEs moyennes sont 1.59 mm pour la similarité basée sur le champ de potentiel, 0.28 mm pour la similarité basée sur la spécificité d'objet, et 0.18 mm pour la similarité hybride qui combine ces deux similarités. Nos tests ont montré que la similarité de champ potentiel est sensible. La RMSE moyenne des composants tibiaux était beaucoup plus élevée que celle des composants fémoraux parce qu'il y a plus d'artefacts autour du composant tibial que celui fémoral. Les contours des composants tibiaux sont attirés par le champ de potentiel des contours tibial et fibulaire. La similarité de spécificité d'objet était plus stable et précise que celle du champ de potentiel. Mais la similarité hybride basée sur les contours et les régions a donné le meilleur résultat en termes de stabilité et de précision. L'avantage de la similarité basée sur les régions est la stabilité et la robustesse contre le bruit. Cependant, cette mesure seule n'est pas précise car le nombre de pixels sur le contour de la région est inférieur au nombre de pixels à l'intérieur de la région. Au contraire, la mesure de similarité basée sur les contours est précise tout en étant plus sensible au bruit ou à d'autres artefacts existant dans les images. C'est pourquoi la combinaison de la stabilité et de la robustesse au bruit de la similarité basée sur les régions avec la précision de la similarité basée sur les contours fournit une méthode de recalage non supervisée robuste.

La méthode de reconstruction 3D des membres inférieurs (rotule, astragale et bassin) à partir de deux images radiographiques biplanaires obliques. Cette approche exploite efficacement toutes les informations existantes de l'image (contours et ROIs) et l'estimation de FastMap à la fois pour la réduction de la dimensionnalité et pour contraindre le problème de reconstruction 3D. Tandis que la méthode biplanaire proposée dans [**18**] n'utilise pas toutes les informations dans deux images (juste les contours) comme notre vraisemblance. De plus, sa réduction de dimensionnalité PCA est linéaire ce qui, en termes de déformations statistiques admissibles, est une hypothèse forte qui n'est pas nécessairement vraie dans notre contexte. Le problème de reconstruction 3D est considéré comme un problème d'optimisation de fonction d'énergie englobant les informations extraites de l'image et ceux de la base de données par le FastMap. Les avantages de notre méthode de reconstruction 3D sont la minimisation de la dose de rayons X (seulement deux images radiographiques obliques), l'exploitation maximale d'informations existant dans ces deux images (contours et ROIs) et l'utilisation de la méthode statistique non linéaire FastMap qui donne une bonne représentation des déformations dans la base de données osseuses par rapport à celles utilisées

dans d'autres méthodes de reconstruction 3D. Les résultats obtenus dans nos expériences sont assez encourageants et indiquent qu'une reconstruction 3D précise et non supervisée est réalisable. Cette approche a été validée sur un ensemble de 13 os (5 rotules, 5 astragales et 3 bassins) de 13 patients (13 paires d'images radiographiques des membres inférieurs), en comparant le modèle obtenu de notre approche et celles obtenues avec CT-Scan, qui a été considérée comme la vérité terrain. La RMS moyenne était respectivement 0,9 mm, 1,4 mm et 3,4 mm pour la rotule, l'astragale et le bassin. La moyenne et l'écart type des pourcentages pour les points dont l'erreur est inférieure à 1 mm de la rotule et de l'astragale étaient $75{,}1 \pm 10{,}1\%$ et $51{,}6 \pm 5{,}1\%$. Ceux dont l'erreur est inférieure à 2 mm du bassin étaient $52{,}7 \pm 9{,}8\%$. Les expérimentations ont montré que nos résultats sont meilleurs que ceux de la méthode de modélisation statistique hiérarchique [**18**] dans le cas de la rotule et de l'astragale. Par contre, la précision de la reconstruction 3D du bassin était moindre que celle des autres os car le bassin a une structure très complexe et sa base de données était petite, limitée à 39 objets dans notre application. De plus, le sacrum et le coccyx étaient flous sur les images en plus du fait qu'il s'avère difficile de distinguer le sacrum de la colonne vertébrale.

Les principales contributions de cette thèse sont : d'abord, la construction et l'utilisation de la base d'apprentissage multi-atlas qui nous a permis d'estimer, de manière non paramétrique, une collection de superpixels qui capturent toute la variabilité non linéaire et locale des ROIs présents dans la base. Cela nous a permis à la fois d'augmenter la précision de l'étape d'élagage de la carte superpixel tout en simplifiant le problème d'optimisation impliqué dans cette étape cruciale. Deuxièmement, la fusion par propagation d'étiquettes est réalisée dans le sens de la variation de l'information et nous a permis de raffiner le contour externe de la structure osseuse à segmenter, et surtout de déduire les étiquettes de région interne à partir des segmentations présélectionnées de l'atlas au résultat final de la segmentation. Troisièmement, une mesure de similarité hybride basée sur les contours et les régions / ROIs permet d'améliorer la stabilisation et la précision des approches grâce à la robustesse contre les bruits et la stabilisation de la similarité basée sur les régions et la précision de la similarité basée sur les contours. Quatrièmement, l'utilisation de deux images radiographiques biplanaires obliques permet d'exploiter toutes les informations dans deux images parce que les structures osseuses dans les deux images obliques apparaissent moins claires que dans celle PA, mais beaucoup plus claires que dans celle LAT. Cinquièmement, une nouvelle approche simple et non supervisée de positionner les os dans l'environnement radiographique est présentée. Elle calcule au préalable un environnement de référence contenant tous les os à reconstruire. Pour chaque paire d'images en entrée, elle crée un environnement puis le recale sur celui de référence. Ce positionnement peut être ensuite raffiné par un recalage rigide 3D/2D. Sixièmement, la méthode statistique non linéaire FastMap est utilisée pour

créer un SDM pour la reconstruction 3D. Elle permet de fournir une bonne représentation des déformations statistiques admissibles dans la base de données osseuses.

Les approches proposées dans cette thèse offrent des possibilités intéressantes pour divers problèmes médicaux nécessitant une connaissance de l'information 3D. Le recalage 3D / 2D retourne les positions des implants fémoraux et tibiaux dans l'environnement radiographique. Elles permettent de mesurer et analyser ces positions et ces orientations afin d'étudier l'alignement de ces implants sur les trois plans (sagittal, frontal et transversal). Par exemple, la rotation externe des implants peut être mesurée. Ces rotations sont importantes dans le suivi fémoro-patellaire car, une rotation inappropriée du composant fémoral peut entraîner un déséquilibre en flexion et des problèmes fémoro-patellaires [**83**]. Ces paramètres morphologiques sont souvent calculés soit sur le plan sagittal ou frontal via les radiographies, soit sur le plan transversal via le CT-scan. Ces positions en 3D conduisent la possibilité de calculer les nouveaux paramètres. De plus, notre méthode non supervisée permet d'effectuer facilement et précisément ces analyses sur un grand nombre de patients. Les modèles 3D de la rotule, de l'astragale et du bassin produits par la reconstruction 3D servent la simulation des traitements orthopédiques. Le principal avantage d'un simulateur est l'accès à de l'information impossible à obtenir seulement par l'observation. Grâce à ces modèles 3D, les mesures chirurgicales pertinentes peuvent être calculées. La rotule est pour l'évaluation du mécanisme extenseur du genou. L'astragale est pour l'analyse de l'alignement entre la cheville et du pied. Le bassin est pour l'évaluation l'alignement et l'orientation de la hanche et la colonne vertébrale. L'utilisation du bassin personnalisé a le potentiel de fournir un meilleur repère pour les systèmes de navigation chirurgicale. De plus, en combinant avec la reconstruction des os longs tels que le fémur, tibia et péroné, un squelette complet des membres inférieurs (du bassin à l'astragale) peut être créé à partir juste deux images radiographiques biplanaires. Il permet de commencer des nouvelles études nécessitant un squelette complet des membres inférieurs pour en connaître l'alignement et l'orientation relative des os.

Conclusion générale

L'objectif de cette thèse était à développer et valider une méthode de reconstruction 3D des membres inférieurs (rotule, astragale et bassin) à partir de deux images radiographiques basses-doses calibrées (vues en 45 et 135 degrées).

Cette thèse a commencé par expliquer les inconvénients de CT-Scan et MRI afin d'introduire les avantages intéressants de la radiographie. Cependant la radiographie a aussi ses limites qui sont la variété du bruit d'imagerie et le chevauchement entre les os et les tissus mous denses. C'est pourquoi nous avons proposé une méthode de segmentation des images radiographiques basée sur la fusion entropique afin d'extraire les régions osseuses de la rotule, de l'astragale et du bassin. Et une méthode de recalage 3D / 2D était développé pour aider de positionner préliminairement le modèle paramétrique de la reconstruction 3D dans l'environnement radiographique de basse-dose. Nous avons présenté ensuite une revue de littérature sur les différentes techniques de segmentation, de recalage 3D / 2D et de la reconstruction 3D des os à partir de radiographies biplanaires afin de nous permettre d'éclairer les limites existantes et de définir les objectifs de notre projet de recherche.

Dans le cadre de cette thèse, nous avons développé en premier lieu une méthode originale de segmentation des images radiographiques permettant d'extraire des ROIs de la rotule, de l'astragale et du bassin. Cette méthode utilise l'apprentissage d'une base de données des images radiographiques de ces os pré-segmentées et recalées pour construire une carte de superpixels permettant de tenir compte de toute la variabilité locale et non linéaire existante dans la base, puis la propagation d'étiquettes basée sur le concept d'entropie pour raffiner la carte des segmentations en régions internes afin d'obtenir le résultat final. Nous avons aussi développé une méthode de recalage 3D/2D des composants tibiaux et fémoraux de la prothèse du genou sur deux images radiographiques biplanaires calibrées. Cette méthode utilise une mesure de similarité hybride permettant de profiter les avantages et limiter les inconvénients des contours et régions. Nous avons développé une méthode originale de reconstruction 3D non supervisée de la rotule, de l'astragale et du bassin à partir de deux images radiographiques biplanaires oblique calibrées. Cette méthode utilise une connaissance géométrique *a priori*, définie par l'algorithme de réduction de dimensionnalité FastMap, et une vraisemblance basée sur les contours et ROIs. Chaque méthode a ensuite été validée sur

une base de 31 images pour la segmentation, de 64 composants pour le recalage, et de 13 os pour la reconstruction 3D.

Notre approche fournit une représentation précise de la rotule, de l'astragale et du bassin à partir de seulement deux vues radiographiques, tandis que le CT-Scan exige des centaines d'images pour obtenir la même reconstruction 3D. La méthode proposée est donc intéressante en termes de quantité de données à acquérir, traiter et gérer. En plus, deux images radiographiques sont moins coûteuses qu'un scan et le patient subit moins d'irradiation pour un examen. Ce sont des avantages importants, surtout pour les patients qui nécessitent de passer fréquemment des radiographies.

Notre méthode constitue donc une alternative à la reconstruction 3D de CT-Scan avec des avantages non négligeable et sera d'un grand intérêt pour le diagnostic des déformations de la structure osseuse, la simulation des traitements orthopédiques, etc. Cependant, pour le moment, cette méthode de reconstruction sans amélioration n'est pas adaptée pour les applications de navigation chirurgicale par rapport aux erreurs de reconstruction par CT-Scan de ± 1 mm, le gabarit pour ces applications. À notre connaissance, l'approche proposée était la première méthode de reconstruction 3D pour l'astragale et elle reste généralement suffisante pour être appliquée à d'autres problèmes médicaux de reconstruction pour lesquels une base de données de la structure anatomique est disponible (avec deux vues radiographiques).

Toutefois, des améliorations peuvent être suggérées pour la continuité de ce travail de recherche :

- Augmenter la taille de la base d'apprentissage de chaque os pour la segmentation;
- Augmenter la taille de la base d'apprentissage du bassin pour la reconstruction 3D;
- Étendre la méthode de reconstruction 3D sur les autres structures osseuses dont une base d'apprentissage est disponible;
- Le but de notre segmentation est de fournir des images des ROIs des os pour la reconstruction 3D biplanaire. Alors au lieu de segmenter individuellement les images, on peut segmenter simultanément deux images biplanaires pour profiter des caractéristiques communes des os dans les deux images, telles que la hauteur de l'os, la position verticale de l'os dans les deux images, etc.;
- Pour les structures osseuses complexes comme le bassin, on pourrait les découper en sous-parties puis faire reconstruire chaque partie. Alors la reconstruction 3D des os complexes devient la combinaison des reconstructions 3D des parties de cet os. Par exemple pour le bassin, on peut séparer le sacrum et le coccyx du bassin. Puis on reconstruit indépendamment le sacrum et le coccyx avec le reste du bassin.

Références bibliographiques

[1] YI Abdel-Aziz et HM Karara : Direct linear transformation from comparator co, ordinates into object space coordinates in close range photogrammetry. In ASP Symposium on close range photogrammetry, Fall Church, 1971.

[2] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua et Sabine Süsstrunk : Slic superpixels. Rapport technique, 2010.

[3] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua et Sabine Süsstrunk : Slic superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence, 34(11):2274–2282, 2012.

[4] Emmanuel Ahishakiye, Martin Bastiaan Van Gijzen, Julius Tumwiine, Ruth Wario et Johnes Obungoloch : A survey on deep learning in medical image reconstruction. Intelligent Medicine, 2021.

[5] Paul Aljabar, Rolf A Heckemann, Alexander Hammers, Joseph V Hajnal et Daniel Rueckert : Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. NeuroImage, 46(3):726–738, 2009.

[6] Bernard André, Jean Dansereau et Hubert Labelle : Optimized vertical stereo base radiographic setup for the clinical three-dimensional reconstruction of the human spine. Journal of biomechanics, 27(8):1023–1035, 1994.

[7] Alexander Andreopoulos et John K Tsotsos : Efficient and generalizable statistical models of shape and appearance for analysis of cardiac mri. Medical Image Analysis, 12(3):335–357, 2008.

[8] Xabier Artaechevarria, Arrate Munoz-Barrutia et Carlos Ortiz-de Solórzano : Combination strategies in multi-atlas image segmentation: Application to brain MR data. IEEE Transactions on medical imaging, 28(8):1266–1277, 2009.

[9] Benjamin Aubert, Carlos Vazquez, Thierry Cresson, Stefan Parent et Jacques A de Guise : Toward automated 3d spine reconstruction from biplanar radiographs using cnn for statistical spine model fitting. IEEE Transactions on Medical Imaging, 38(12):2796–2806, 2019.

[10] C É Aubin, Jean Dansereau, F Parent, Hubert Labelle et Jacques A de Guise : Morphometric evaluations of personalised 3d reconstructions and geometric models of the human spine. Medical and biological engineering and computing, 35(6):611–618, 1997.

[11] Wenjia Bai, Wenzhe Shi, Declan P O'Regan, Tong Tong, Haiyan Wang, Shahnaz Jamil-Copley, Nicholas S Peters et Daniel Rueckert : A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: application to cardiac mr images. IEEE transactions on medical imaging, 32(7):1302–1315, 2013.

[12] Nora Baka, Bart L Kaptein, Marleen de Bruijne, Theo van Walsum, JE Giphart, Wiro J Niessen et Boudewijn PF Lelieveldt : 2d–3d shape reconstruction of the distal femur from stereo x-ray imaging using statistical shape models. Medical image analysis, 15(6):840–850, 2011.

[13] E. Bardinet, L.D. Cohen et N. Ayache : A parametric deformable model to fit unstructured 3D data. Computer Vision and Image Understanding, 71(1):39–54, 1998.

[14] William A Barrett et Eric N Mortensen : Interactive live-wire boundary extraction. Medical image analysis, 1(4):331–341, 1997.

[15] Gert Behiels, Dirk Vandermeulen, Frederik Maes, Paul Suetens et Piet Dewaele : Active shape model-based segmentation of digital x-ray images. In Medical Image Computing and Computer-Assisted Intervention, pages 128–137. Springer, 1999.

[16] Lilia Chorfi Belhadj et Max Mignotte : Spatio-temporal fastmap-based mapping for human action recognition. In 2016 IEEE International Conference on Image Processing (ICIP), pages 3046–3050. IEEE, 2016.

[17] Said Benameur, Max Mignotte, François Destrempes et Jacques A De Guise : Three-dimensional biplanar reconstruction of scoliotic rib cage using the estimation of a mixture of probabilistic prior models. IEEE Transactions on Biomedical Engineering, 52(10):1713–1728, 2005.

[18] Said Benameur, Max Mignotte, Hubert Labelle et Jacques A De Guise : A hierarchical statistical modeling approach for the unsupervised 3-d biplanar reconstruction of the scoliotic spine. IEEE Transactions on Biomedical Engineering, 52(12):2041–2057, 2005.

[19] Said Benameur, Max Mignotte, Stefan Parent, Hubert Labelle, Wafa Skalli et Jacques de Guise : 3d/2d registration and segmentation of scoliotic vertebrae using statistical models. Computerized Medical Imaging and Graphics, 27(5):321–337, 2003.

[20] Said Benameur, Max Mignotte, Stefan Parent, Hubert Labelle, Wafa Skalli et Jacques de Guise : 3d/2d registration and segmentation of scoliotic vertebrae using statistical models. Computerized Medical Imaging and Graphics, 27(5):321–337, 2003.

[21] J Benjamin : Component alignment in total knee arthroplasty. Instructional Course Lectures, 55:405–412, 2006.

[22] R. Benjamin : Object-based 3D X-ray imaging. In Computer Vision, Virtual Reality and Robotics in Medicine, pages 444–448, Paris, France, 1995.

[23] P Bertolino et A Montanvert : Coopération régions contours multirésolution en segmentation d'image. Actes du 10e Congrès AFCET/Reconnaissance des formes et intelligence artificielle, Rennes, pages 16–18, 1996.

[24] S Bertrand, S Laporte, S Parent, W Skalli et D Mitton : Three-dimensional reconstruction of the rib cage from biplanar radiography. IRBM, 29(4):278–286, 2008.

[25] Hamidullah Binol, M Khalid Khan Niazi, Alisha Plotner, Jennifer Sopkovich, Benjamin H Kaffenberger et Metin N Gurcan : A multidimensional scaling and sample clustering to obtain a representative subset of training data for transfer learning-based rosacea lesion identification. In Medical Imaging 2020: Computer-Aided Diagnosis, volume 11314, page 1131415. International Society for Optics and Photonics, 2020.

[26] Ingwer Borg et Patrick JF Groenen : Modern multidimensional scaling: Theory and applications. Springer Science & Business Media, 2005.

[27] Antoni Buades, Bartomeu Coll et Jean-Michel Morel : Nonlocal image and movie denoising. International journal of computer vision, 76(2):123–139, 2008.

[28] Antoni Buades, Bartomeu Coll et Jean-Michel Morel : Non-local means denoising. Image Processing On Line, 1:208–212, 2011.

[29] J Callaghan, A Dennis, W Paprosky et A Rosenberg : Patellofemoral complications in total knee arthroplasty. Hip and Knee Reconstruction. Orthopaedic Knowledge Update. American Academy of Orthopaedic Surgeons, 34:283–9, 1995.

[30] John Canny : A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, (6):679–698, 1986.

[31] Xiaohui Cao, Ming Liu, Fushan Zhai, Nan Li, Feng Li, Chaoen Bao, Yinliang Liu et Gang Chen : Comparative evaluation of image registration methods with different interest regions in lung cancer radiotherapy. BMC medical imaging, 19(1):1–6, 2019.

[32] Xuyu Cao, Chen Chen et Lixia Tian : Supervised multidimensional scaling and its application in mri-based individual age predictions. Neuroinformatics, 19(2):219–231, 2021.

[33] Jana Čavojská, Julian Petrasch, Denny Mattern, Nicolas Jens Lehmann, Agnès Voisard et Peter Böttcher : Estimating and abstracting the 3d structure of feline bones using neural networks on x-ray (2d) images. Communications biology, 3(1):1–13, 2020.

[34] Pietro Cerveri, Costanza Sacco, Gianluca Olgiati, Alfonso Manzotti et Guido Baroni : 2d/3d reconstruction of the distal femur using statistical shape models addressing personalized surgical instruments in knee arthroplasty: a feasibility analysis. The International Journal of Medical Robotics and Computer Assisted Surgery, 13(4):e1823, 2017.

[35] Y Chaibi, T Cresson, B Aubert, J Hausselle, P Neyret, O Hauger, JA De Guise et W Skalli : Fast 3d reconstruction of the lower limb using a parametric model and statistical inferences and clinical measurements calculation from biplanar x-rays. Computer methods in biomechanics and biomedical engineering, 15(5):457–466, 2012.

[36] Hui Chen et Bir Bhanu : Efficient recognition of highly similar 3d objects in range images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(1):172–179, 2008.

[37] Ying Chen, Xianhe Ee, Wee Kheng Leow et Tet Sen Howe : Automatic extraction of femur contours from hip x-ray images. In International Workshop on Computer Vision for Biomedical Image Applications, pages 200–209. Springer, 2005.

[38] Timothy F Cootes, Andrew Hill, Christopher J Taylor et Jane Haslam : The use of active shape models for locating structures in medical images. In Biennial International Conference on Information Processing in Medical Imaging, pages 33–47. Springer, 1993.

[39] Pierrick Coupé, José V Manjón, Vladimir Fonov, Jens Pruessner, Montserrat Robles et D Louis Collins : Nonlocal patch-based label fusion for hippocampus segmentation. In 13th International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010, volume 6363, pages 129–136. Springer, September 20-24 2010.

[40] Michael AA Cox et Trevor F Cox : Multidimensional scaling. In Handbook of data visualization, pages 315–347. Springer, 2008.

[41] Trevor F Cox et Michael AA Cox : Multidimensional scaling. CRC press, 2000.

[42] X Cufi, X Munoz, J Freixenet et J Marti : A review of image segmentation techniques integrating region and boundary information. Advances in imaging and electron physics, 120:1–39, 2003.

[43] Samuel Dambreville, Yogesh Rathi et Allen Tannenbaum : A framework for image segmentation using shape models and kernel space shape priors. IEEE transactions on pattern analysis and machine intelligence, 30(8):1385–1399, 2008.

[44] Jean Dansereau et Ian AF Stokes : Measurements of the three-dimensional shape of the rib cage. Journal of biomechanics, 21(11):893–901, 1988.

[45] Francois Destrempes, Max Mignotte et Jean-Francois Angers : Localization of shapes using statistical models and stochastic optimization. IEEE Transactions on pattern analysis and machine intelligence, 29(9):1603–1615, 2007.

[46] Francois Destrempes, Max Mignotte et Jean-Francois Angers : Localization of shapes using statistical models and stochastic optimization. IEEE transactions on pattern analysis and machine intelligence, 29(9):1603–1615, 2007.

[47] Sonja Dieterich, Carlo Cavedon, Cynthia F Chuang, Alan B Cohen, Jeffrey A Garrett, Charles L Lee, Jessica R Lowenstein, Maximian F d'Souza, David D Taylor Jr, Xiaodong Wu et al. : Report of aapm tg 135: quality assurance for robotic radiosurgery. Medical physics, 38(6Part1):2914–2936, 2011.

[48] Feng Ding, Wee Leow et Tet Howe : Automatic segmentation of femur bones in anterior-posterior pelvis x-ray images. In Computer Analysis of Images and Patterns, pages 205–212. Springer, 2007.

[49] Jason A Dowling, Jurgen Fripp, Shekhar Chandra, Josien P W Pluim, Jonathan Lambert, Joel Parker, James Denham, Peter B Greer et Olivier Salvado : Fast automatic multi-atlas segmentation of the prostate from 3D MR images. In Proceedings of the 2011 International Conference on Prostate Cancer Imaging: Image Analysis and Image-guided Interventions, MICCAI'11, pages 10–21, Berlin, Heidelberg, 2011. Springer-Verlag.

[50] Jamil Draréni, Sébastien Roy et Peter Sturm : Plane-based calibration for linear cameras. International Journal of Computer Vision, 91(2):146–156, 2011.

[51] Jalda Dworzak, Hans Lamecker, Jens von Berg, Tobias Klinder, Cristian Lorenz, Dagmar Kainmüller, Heiko Seim, Hans-Christian Hege et Stefan Zachow : 3d reconstruction of the human rib cage from 2d projection images using a statistical shape model. International journal of computer assisted radiology and surgery, 5(2):111–124, 2010.

[52] Moritz Ehlke, Heiko Ramm, Hans Lamecker, Hans-Christian Hege et Stefan Zachow : Fast generation of virtual x-ray images for reconstruction of 3d anatomy. IEEE transactions on visualization and computer graphics, 19(12):2673–2682, 2013.

[53] Hatem Amin Abdel Fattah El Dakhakhni : Reconstruction of patient-specific bone models from x-ray radiography. 2013.

[54] S Diop El Hadji et Valérie Burdin : Bi-planar image segmentation based on variational geometrical active contours with shape priors. Medical image analysis, 17(2):165–181, 2013.

[55] Christos Faloutsos et King-Ip Lin : Fastmap: A fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. In Proceedings of the 1995 ACM SIGMOD international conference on Management of data, pages 163–174, 1995.

[56] Pedro F Felzenszwalb et Daniel P Huttenlocher : Efficient graph-based image segmentation. International journal of computer vision, 59(2):167–181, 2004.

[57] Markus Fleute et Stéphane Lavallée : Building a complete surface model from sparse data using statistical shape models: Application to computer assisted knee surgery. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 879–887. Springer, 1998.

[58] Markus Fleute et Stéphane Lavallée : Nonrigid 3-d/2-d registration of images using statistical models. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 138–147. Springer, 1999.

[59] Ted Julien Tchinde Fotsin, Carlos Vázquez, Thierry Cresson et Jacques De Guise : Shape, pose and density statistical model for 3d reconstruction of articulated structures from x-ray images. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 2748–2751. IEEE, 2019.

[60] Olivier François et al. : Global optimization with exploration/selection algorithms and simulated annealing. The Annals of Applied Probability, 12(1):248–271, 2002.

[61] PE Galibarov, PJ Prendergast et AB Lennon : A method to reconstruct patient-specific proximal femur surface models from planar pre-operative radiographs. Medical engineering & physics, 32(10): 1180–1188, 2010.

[62] Pavan Gamage, Sheng Quan Xie, Patrice Delmas et Peter Xu : 3d reconstruction of patient specific bone models from 2d radiographs for image guided orthopedic surgery. In 2009 Digital Image Computing: Techniques and Applications, pages 212–216. IEEE, 2009.

[63] Rohit Girdhar, David F Fouhey, Mikel Rodriguez et Abhinav Gupta : Learning a predictable and generative vector representation for objects. In European Conference on Computer Vision, pages 484–499. Springer, 2016.

[64] George R Hanson, Jeremy F Suggs, Andrew A Freiberg, Sridar Durbhakula et Guoan Li : Investigation of in vivo 6dof total knee arthoplasty kinematics using a dual orthogonal fluoroscopic system. Journal of Orthopaedic Research, 24(5):974–981, 2006.

[65] Rola Harmouche, Farida Cheriet, Hubert Labelle et Jean Dansereau : Multimodal image registration of the scoliotic torso for surgical planning. BMC Medical Imaging, 13(1):1–12, 2013.

[66] Phlipp Henzler, Volker Rasche, Timo Ropinski et Tobias Ritschel : Single-image tomography: 3d volumes from 2d cranial x-rays. In Computer Graphics Forum, volume 37, pages 377–388. Wiley Online Library, 2018.

[67] Jeroen Hermans, Johan Bellemans, Frederik Maes, Dirk Vandermeulen et Paul Suetens : A statistical framework for the registration of 3d knee implant components to single-plane x-ray images. In 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pages 1–8. IEEE, 2008.

[68] Paul R Hill, Cedric Nishan Canagarajah et David R Bull : Image segmentation using a texture gradient based watershed transform. IEEE Transactions on Image Processing, 12(12):1618–1633, 2003.

[69] Geoffrey E Hinton et Ruslan R Salakhutdinov : Reducing the dimensionality of data with neural networks. science, 313(5786):504–507, 2006.

[70] Minh A Hoang, Jan-Mark Geusebroek et Arnold WM Smeulders : Color texture measurement and segmentation. Signal processing, 85(2):265–275, 2005.

[71] Mafruha Mowin Hossain, Abdullah A Muhit, Mark R Pickering, Jennie Scarvell et Paul Smith : A 3d-2d image registration algorithm for kinematic analysis of the knee after total knee arthroplasty (tka). In 2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pages 1–6. IEEE, 2013.

[72] Tzung-Chi Huang, Geoffrey Zhang, Thomas Guerrero, George Starkschall, Kan-Ping Lin et Ken Forster : Semi-automated ct segmentation using optic flow and fourier interpolation techniques. Computer methods and programs in biomedicine, 84(2-3):124–134, 2006.

[73] Jie Huo, Guanghui Wang, QM Jonathan Wu et Akilan Thangarajah : Label fusion for multi-atlas segmentation based on majority voting. In International Conference Image Analysis and Recognition, pages 100–106. Springer, 2015.

[74] Nathaniel P Jacobson et Maya R Gupta : Design goals and solutions for display of hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing, 43(11):2684–2692, 2005.

[75] Anil K. Jain, Yu Zhong et Sridhar Lakshmanan : Object matching using deformable templates. IEEE Transactions on pattern analysis and machine intelligence, 18(3):267–278, 1996.

[76] David Antonio Gómez Jáuregui et Patrick Horain : Region-based vs. edge-based registration for 3d motion capture by real time monoscopic vision. In International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications, pages 344–355. Springer, 2009.

[77] Pierre-Marc Jodoin, Max Mignotte et Christophe Rosenberger : Segmentation framework based on label field fusion. IEEE Transactions on image Processing, 16(10):2535–2550, 2007.

[78] Sushitha Susan Joseph et D Aju : A comparative survey on three-dimensional reconstruction of medical modalities based on various approaches. In Information Systems Design and Intelligent Applications, pages 223–233. Springer, 2019.

[79] Davood Karimi, Golnoosh Samei, Claudia Kesch, Guy Nir et Septimiu E Salcudean : Prostate segmentation in mri using a convolutional neural network architecture and training strategy based on statistical shape models. International journal of computer assisted radiology and surgery, 13(8):1211–1219, 2018.

[80] Michael Kass, Andrew Witkin et Demetri Terzopoulos : Snakes: Active contour models. International journal of computer vision, 1(4):321–331, 1988.

[81] Yoni Kasten, Daniel Doktofsky et Ilya Kovler : End-to-end convolutional neural network for 3d reconstruction of knee bones from bi-planar x-ray images. In International Workshop on Machine Learning for Medical Image Reconstruction, pages 123–133. Springer, 2020.

[82] Hangkee Kim, Kisuk Lee, Dongchun Lee et Nakhoon Baek : 3d reconstruction of leg bones from x-ray images using cnn-based feature analysis. In 2019 International Conference on Information and Communication Technology Convergence (ICTC), pages 669–672. IEEE, 2019.

[83] Youngjun Kim, Kang-Il Kim, Jin hyeok Choi et Kunwoo Lee : Novel methods for 3d postoperative analysis of total knee arthroplasty using 2d–3d image registration. Clinical Biomechanics, 26(4):384–391, 2011.

[84] Y. Kita : Elastic-model driven analysis of several views of a deformable cylindrical object. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(12):1150–1162, 1996.

[85] Koichi Kobayashi, Makoto Sakamoto, Yuji Tanabe, Akihiro Ariumi, Takashi Sato, Go Omori et Yoshio Koga : Automated image registration for assessing three-dimensional alignment of entire lower extremity and implant position using bi-plane radiography. Journal of Biomechanics, 42(16):2818–2822, 2009.

[86] William J Kostis, Anthony P Reeves, David F Yankelevitz, Claudia I Henschke et al. : Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical ct images. IEEE Trans. Med. Imaging, 22(10):1259–1274, 2003.

[87] Peter R Krekel, Charl P Botha, Edward R Valstar, Paul W de Bruin, Piet M Rozing et Frits H Post : Interactive simulation and comparative visualisation of the bone-determined range of motion of the human shoulder. In SimVis, pages 275–288, 2006.

[88] Margaret M Landy et Peter S Walker : Wear of ultra-high-molecular-weight polyethylene components of 90 retrieved knee prostheses. The Journal of arthroplasty, 3:S73–S85, 1988.

[89] S Laporte, W Skalli, JA De Guise, F Lavaste et D Mitton : A biplanar reconstruction method based on 2d and 3d contours: application to the distal femur. Computer Methods in Biomechanics & Biomedical Engineering, 6(1):1–6, 2003.

[90] Maxime Latulippe : Calage robuste et accéléré de nuages de points en environnements naturels via l'apprentissage automatique. 2013.

[91] Anthony Le Bras, S Laporte, V Bousson, David Mitton, Jacques A De Guise, JD Laredo et Wafa Skalli : 3d reconstruction of the proximal femur with low-dose digital stereoradiography. Computer Aided Surgery, 9(3):51–57, 2004.

[92] Fang Liu, Zhaoye Zhou, Hyungseok Jang, Alexey Samsonov, Gengyan Zhao et Richard Kijowski : Deep convolutional neural network and 3d deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magnetic resonance in medicine, 79(4):2379–2391, 2018.

[93] Cristian Lorenz et Nils Krahnstöver : Generation of point-based 3d statistical shape models for anatomical objects. Computer vision and image understanding, 77(2):175–191, 2000.

[94] Jyrki Lötjönen, Pierre-Jean Reissman, Isabelle E Magnin et Toivo Katila : Model extraction from magnetic resonance volume data using the deformable pyramid. Medical Image Analysis, 3(4):387–406, 1999.

[95] Laurens van der Maaten et Geoffrey Hinton : Visualizing data using t-sne. Journal of Machine Learning Research, 9(Nov):2579–2605, 2008.

[96] Mohamed R Mahfouz, William A Hoff, Richard D Komistek et Douglas A Dennis : A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images. IEEE transactions on medical imaging, 22(12):1561–1574, 2003.

[97] Primoz Markelj, Dejan Tomaževič, Bostjan Likar et Franjo Pernuš : A review of 3d/2d registration methods for image-guided interventions. Medical image analysis, 16(3):642–661, 2012.

[98] David R Martin, Charless C Fowlkes et Jitendra Malik : Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE transactions on pattern analysis and machine intelligence, 26(5):530–549, 2004.

[99] Genaro Tolentino Marzan : Rational design for close-range photogrammetry. PhD thesis, 1976.

[100] Kevin R Math, Syed Furqan Zaidi, Catherine Petchprapa et Steven F Harwin : Imaging of total knee arthroplasty. In Seminars in musculoskeletal radiology, volume 10, pages 047–063. Copyright© 2006 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ..., 2006.

[101] Shuichi Matsuda, H Miura, Ryuji Nagamine, Ken Urabe, Go Hirata et Yukihide Iwamoto : Effect of femoral and tibial component position on patellar tracking following total knee arthroplasty: 10-year follow-up of miller-galante i knees. The American journal of knee surgery, 14(3):152–156, 2001.

[102] Georg Matziolis, Doerte Krocker, Ulrike Weiss, Stephan Tohtz et Carsten Perka : A prospective, randomized study of computer-assisted and conventional total knee arthroplasty: three-dimensional evaluation of implant alignment and rotation. JBJS, 89(2):236–243, 2007.

[103] Marina Meilă : Comparing clusterings: an axiomatic view. In Proceedings of the 22nd International conference on Machine learning, pages 577–584, Bonn, Germany, 2005. ACM.

[104] Marina Meilă : Comparing clusterings-an information based distance. Journal of multivariate analysis, 98(5):873–895, 2007.

[105] Max Mignotte : A de-texturing and spatially constrained K-means approach for image segmentation. Pattern Recognition Letters, 32(2):359–367, 2011.

[106] Max Mignotte : MDS-based multiresolution nonlinear dimensionality reduction model for color image segmentation. IEEE Transactions on neural networks, 22(3):447–460, 2011.

[107] Max Mignotte : Mds-based multiresolution nonlinear dimensionality reduction model for color image segmentation. IEEE transactions on neural networks, 22(3):447–460, 2011.

[108] Max Mignotte : Mds-based segmentation model for the fusion of contour and texture cues in natural images. Computer Vision and Image Understanding, 116(9):981–990, 2012.

[109] Max Mignotte : A label field fusion model with a variation of information estimator for image segmentation. Information Fusion, 20:7–20, 2014.

[110] Max Mignotte : A non-stationary MRf model for image segmentation from a soft boundary map. Pattern Analysis and Applications, 17(1):129–139, 2014.

[111] Max Mignotte, Christophe Collet, Patrick Perez et Patrick Bouthemy : Sonar image segmentation using an unsupervised hierarchical MRF model. IEEE Transactions on image processing, 9(7):1216–1231, 2000.

[112] Max Mignotte et Jean Meunier : A multiscale optimization approach for the dynamic contour-based boundary detection issue. Computerized Medical Imaging and Graphics, 25(3):265–275, 2001.

[113] Max Mignotte, Jean Meunier et Jean-Paul Soucy : Dct-based complexity regularization for em tomographic reconstruction. IEEE Transactions on Biomedical Engineering, 55(2):801–805, 2008.

[114] Max Mignotte, Jean Meunier et Jean-Claude Tardif : Endocardial boundary e timation and tracking in echocardiographic images using deformable template and markov random fields. Pattern Analysis & Applications, 4(4):256–271, 2001.

[115] Max Mignotte, Jean Meunier et Jean-Claude Tardif : Endocardial boundary e timation and tracking in echocardiographic images using deformable template and markov random fields. Pattern Analysis & Applications, 4(4):256–271, 2001.

[116] D Mitton, C Landry, S Veron, Wata Skalli, F Lavaste et Jacques A De Guise : 3d reconstruction method from biplanar radiography using non-stereocorresponding points and elastic deformable meshes. Medical and Biological Engineering and Computing, 38(2):133–139, 2000.

[117] David Mitton, S Deschenes, S Laporte, B Godbout, S Bertrand, Jacques A de Guise et W Skalli : 3d reconstruction of the pelvis from bi-planar radiography. Computer methods in biomechanics and biomedical engineering, 9(1):1–5, 2006.

[118] Jean-Philippe Morin, Christian Desrosiers et Luc Duong : A random walk approach for multiatlas-based segmentation. In Pattern Recognition (ICPR), 2012 21st International Conference on, pages 3636–3639. IEEE, 2012.

[119] Daniel C Moura, Jonathan Boisvert, Jorge G Barbosa, Hubert Labelle et João Manuel RS Tavares : Fast 3d reconstruction of the spine from biplanar radiographs using a deformable articulated model. Medical engineering & physics, 33(8):924–933, 2011.

[120] Xavier Muñoz, Jordi Freixenet, X Cufi et J Marti : Strategies for image segmentation combining region and boundary information. Pattern recognition letters, 24(1):375–392, 2003.

[121] Andriy Myronenko et Xubo Song : Point set registration: Coherent point drift. IEEE transactions on pattern analysis and machine intelligence, 32(12):2262–2275, 2010.

[122] Dac Cong Tai Nguyen, Said Benameur, Max Mignotte et Frédéric Lavoie : Superpixel and entropy-based multi-atlas fusion framework for the segmentation of x-ray images. In 18th International Conference on Image Analysis and Processing, ICIAP'15, Lecture Notes in Computer Science, volume 9280, pages 151–161, Genova, Italy, September 2015.

[123] Dac Cong Tai Nguyen, Said Benameur, Max Mignotte et Frédéric Lavoie : Superpixel and multi-atlas based fusion entropic model for the segmentation of x-ray images. Medical Image Analysis, 48:58–74, 2018.

[124] B. Nikkhade-Dehkordi, M. Bro-Nielsen, T. Darvann, C. Gramkow, N. Egund et K. Hermann : 3D reconstruction of the femoral bone using two X-ray images from orthogonal views. In Computer Assisted Radiology, page 1015, June 26-29 1996.

[125] T Ooura : General purpose FFT (fast fourier/cosine/sine transform) package. http://momonga.t.u-tokyo.ac.jp/∼ooura/fft.html.

[126] Mark J Pearcy : Stereo radiography of lumbar spine motion. Acta Orthopaedica Scandinavica, 56(sup212):1–45, 1985.

[127] Carlos Andrés Pena-Reyes et Moshe Sipper : Evolutionary computation in medicine: an overview. Artificial Intelligence in Medicine, 19(1):1–23, 2000.

[128] Terry M Peters : Image-guidance for surgical procedures. Physics in Medicine & Biology, 51(14):R505, 2006.

[129] S Quijano, A Serrurier, B Aubert, S Laporte, P Thoreux et W Skalli : Three-dimensional reconstruction of the lower limb from biplanar calibrated radiographs. Medical engineering & physics, 35(12):1703–1712, 2013.

[130] Xiaofeng Ren et Jitendra Malik : Learning a classification model for segmentation. In 9th IEEE International Conference on Computer Vision, volume 1, page 10. IEEE, 2003.

[131] Cornelius Johannes Frederik Reyneke, Marcel Lüthi, Valérie Burdin, Tania S Douglas, Thomas Vetter et Tinashe EM Mutsvangwa : Review of 2-d/3-d reconstruction using statistical shape and intensity models and x-ray image synthesis: Toward a unified framework. IEEE reviews in biomedical engineering, 12:269–286, 2018.

[132] Seung Joon Rhee, Jae Young Cho, Seung Hyeon Jeung, Kein Boon Poon, Yoon Young Choi et Jeung Tak Suh : Combined rotational alignment change after total knee arthroplasty in different tibial component designs: implications for optimal tibial component rotational alignment. Knee surgery & related research, 30(1):74, 2018.

[133] Torsten Rohlfing, Robert Brandt, Randolf Menzel et Calvin R Maurer : Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage, 21(4):1428–1442, 2004.

[134] Torsten Rohlfing et Calvin R Maurer : Shape-based averaging. IEEE Transactions on Image Processing, 16(1):153–161, 2007.

[135] Torsten Rohlfing, Daniel B Russakoff et Calvin R Maurer : Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. IEEE Transactions on medical imaging, 23(8):983–994, 2004.

[136] Françcois Rousseau, Piotr A Habas et Colin Studholme : A supervised patch-based approach for human brain labeling. IEEE Transactions on medical imaging, 30(10):1852–1862, 2011.

[137] Sam T Roweis et Lawrence K Saul : Nonlinear dimensionality reduction by locally linear embedding. science, 290(5500):2323–2326, 2000.

[138] Szymon Rusinkiewicz et Marc Levoy : Efficient variants of the icp algorithm. In Proceedings third international conference on 3-D digital imaging and modeling, pages 145–152. IEEE, 2001.

[139] Sanjeev Sabharwal et Caixia Zhao : Assessment of lower limb alignment: supine fluoroscopy compared with a standing full-length radiograph. JBJS, 90(1):43–51, 2008.

[140] M Saleh, P Harriman et DJ Edwards : A radiological method for producing precise limb alignment. The Journal of bone and joint surgery. British volume, 73(3):515–516, 1991.

[141] Takashi Sato, Yoshio Koga et Go Omori : Three-dimensional lower extremity alignment assessment system: application to evaluation of component position after total knee arthroplasty. The Journal of arthroplasty, 19(5):620–628, 2004.

[142] Takashi Sato, Yoshio Koga, Ten Sobue, Go Omori, Yuji Tanabe et Makoto Sakamoto : Quantitative 3-dimensional analysis of preoperative and postoperative joint lines in total knee arthroplasty: a new concept for evaluation of component alignment. The Journal of arthroplasty, 22(4):560–568, 2007.

[143] Jérôme Schmid, Jinman Kim et Nadia Magnenat-Thalmann : Robust statistical shape models for mri bone segmentation in presence of small field of view. Medical image analysis, 15(1):155–168, 2011.

[144] Heiko Seim, Dagmar Kainmueller, Markus Heller, Hans Lamecker, Stefan Zachow et Hans-Christian Hege : Automatic segmentation of the pelvic bones from ct data based on a statistical shape model. VCBM, 8:93–100, 2008.

[145] James A Sethian : Level set methods, evolving interfaces in geometry, fluid mechanics comuputer vision, and materials sciences. Cambridge Monographs on Applied and Computational Mathematics, 3, 1996.

[146] Wenzhe Shi, Herve Lombaert, Wenjia Bai, Christian Ledig, Xiahai Zhuang, Antonio Marvao, Timothy Dawes, Declan O'Regan et Daniel Rueckert : Multi-atlas spectral patchmatch: application to cardiac image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI, volume 8673, pages 348–355. Springer, 2014.

[147] Ryoya Shiode, Mototaka Kabashima, Yuta Hiasa, Kunihiro Oka, Tsuyoshi Murase, Yoshinobu Sato et Yoshito Otake : 2d–3d reconstruction of distal forearm bone from actual x-ray images of the wrist using convolutional neural networks. Scientific Reports, 11(1):1–12, 2021.

[148] Robert Singerman, Hector D Pagan, Amos B Peyser et Victor M Goldberg : Effect of femoral component rotation and patellar design on patellar forces. Clinical orthopaedics and related research, (334):345–353, 1997.

[149] Robert A Siston, Nicholas J Giori, Stuart B Goodman et Scott L Delp : Surgical navigation for total knee arthroplasty: a perspective. Journal of biomechanics, 40(4):728–735, 2007.

[150] Rebecca Smith, Kayvan Najarian et Kevin Ward : A hierarchical method based on active shape models and directed hough transform for segmentation of noisy biomedical images; application in segmentation of pelvic x-ray images. BMC medical informatics and decision making, 9(1):S2, 2009.

[151] Luc Soler, Hervé Delingette, Grégoire Malandain, Johan Montagnat, Nicholas Ayache, Christophe Koehl, Olivier Dourthe, Benoit Malassagne, Michelle Smith, Didier Mutter et al. : Fully automatic anatomical, pathological, and functional segmentation from ct scans for hepatic surgery. Computer Aided Surgery, 6(3):131–142, 2001.

[152] Hideyuki Suenaga, Huy Hoang Tran, Hongen Liao, Ken Masamune, Takeyoshi Dohi, Kazuto Hoshi et Tsuyoshi Takato : Vision-based markerless registration using stereo vision and an augmented reality surgical navigation system: a pilot study. BMC medical imaging, 15(1):1–11, 2015.

[153] Jinwei Sun, Zhiyong Chi et Shaobin Sun : 3d reconstruction of patient-specific femurs using coherent point drift. 2013.

[154] Shaobin Sun, Bin Zhang, Shang Meng, Dan Liu et Jinwei Sun : An improved interactive segmentation method for extracting the edge features of femur digital radiographs. In 2012 International Workshop on Image Processing and Optical Engineering, pages 83351X–83351X. International Society for Optics and Photonics, 2012.

[155] Joshua B Tenenbaum, Vin De Silva et John C Langford : A global geometric framework for nonlinear dimensionality reduction. science, 290(5500):2319–2323, 2000.

[156] D. Terzopoulos, A. Witkin et M. Kass : Constraints on deformable models: Recovering 3D shape and nonrigid motion. Artificial Intelligence, 36(1):91–123, 1988.

[157] Pedro MB Torres, J Miguel Sanches, Paulo JS Gonçalves et Jorge MM Martins : 3d femur reconstruction using a robotized ultrasound probe. In 2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), pages 884–888. IEEE, 2012.

[158] Redha Touati et Max Mignotte : Mds-based multi-axial dimensionality reduction model for human action recognition. In 2014 Canadian Conference on Computer and Robot Vision, pages 262–267. IEEE, 2014.

[159] Redha Touati et Max Mignotte : A multidimensional scaling optimization and fusion approach for the unsupervised change detection problem in remote sensing images. In 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), pages 1–6. IEEE, 2016.

[160] Shubham Tulsiani, Tinghui Zhou, Alexei A Efros et Jitendra Malik : Multi-view supervision for single-view reconstruction via differentiable ray consistency. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2626–2634, 2017.

[161] Edward R Valstar, Rob GHH Nelissen, Johan HC Reiber et Piet M Rozing : The use of roentgen stereophotogrammetry to study micromotion of orthopaedic implants. ISPRS journal of photogrammetry and remote sensing, 56(5-6):376–389, 2002.

[162] Iris Vanhamel, Antonis Katartzis et Hichem Sahli : Hierarchical segmentation via a diffusion scheme in color/texture feature space. In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, volume 1, pages I–969. IEEE, 2003.

[163] Sandro Vega-Pons et José Ruiz-Shulcloper : A survey of clustering ensemble algorithms. International Journal of Pattern Recognition and Artificial Intelligence, 25(03):337–372, 2011.

[164] Edgar A Wakelin, Linda Tran, Joshua G Twiggs, Willy Theodore, Justin P Roe, Michael I Solomon, Brett A Fritsch et Brad P Miles : Accurate determination of post-operative 3d component positioning in total knee arthroplasty: the aurora protocol. Journal of Orthopaedic Surgery and Research, 13, 2018.

[165] Hongzhi Wang, Jung W Suh, Sandhitsu R Das, John B Pluta, Caryne Craige et Paul A Yushkevich : Multi-atlas segmentation with joint label fusion. IEEE transactions on pattern analysis and machine intelligence, 35(3):611–623, 2013.

[166] Hongzhi Wang, Jung Wook Suh, Sandhitsu Das, John Pluta, Murat Altinay et Paul Yushkevich : Regression-based label fusion for multi-atlas segmentation. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pages 1113–1120. IEEE, 2011.

[167] Hongzhi Wang et Paul A Yushkevich : Multi-atlas segmentation without registration: a supervoxel-based approach. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 535–542. Springer, 2013.

[168] Kilian Q Weinberger et Lawrence K Saul : An introduction to nonlinear dimensionality reduction by maximum variance unfolding. In AAAI, volume 6, pages 1683–1686, 2006.

[169] Kilian Q Weinberger, Fei Sha et Lawrence K Saul : Learning a kernel matrix for nonlinear dimensionality reduction. In Proceedings of the twenty-first international conference on Machine learning, page 106. ACM, 2004.

[170] G Wu, Q Wang, D Zhang et D Shen : Robust patch-based multi-atlas labeling by joint sparsity regularization. In MICCAI Workshop on Sparsity Techniques in Medical Imaging (STMI), 2012.

[171] Guorong Wu, Minjeong Kim, Gerard Sanroma, Qian Wang, Brent C Munsell, Dinggang Shen, Alzheimer's Disease Neuroimaging Initiative et al. : Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. NeuroImage, 106:34–46, 2015.

[172] Zhong Xue, Kelvin Wong et Stephen TC Wong : Joint registration and segmentation of serial lung ct images for image-guided lung cancer diagnosis and therapy. Computerized Medical Imaging and Graphics, 34(1):55–60, 2010.

[173] Takaharu Yamazaki, Tetsu Watanabe, Yoshikazu Nakajima, Kazuomi Sugamoto, Tetsuya Tomita, Hideki Yoshikawa et Shinichi Tamura : Improvement of depth position in 2-d/3-d registration of knee implants using single-plane fluoroscopy. IEEE transactions on medical imaging, 23(5):602–612, 2004.

[174] Guoshen Yu et Guillermo Sapiro : Dct image denoising: a simple and effective image denoising algorithm. Image Processing On Line, 1:292–296, 2011.

[175] Ning Yu, Hongzhi Wang et Paul A Yushkevich : Supervoxel-based hierarchical markov random field framework for multi-atlas segmentation. In International Workshop on Patch-based Techniques in Medical Imaging, pages 100–108. Springer, 2016.

[176] Ju Zhang et Thor F Besier : Accuracy of femur reconstruction from sparse geometric data using a statistical shape model. Computer methods in biomechanics and biomedical engineering, 20(5):566–576, 2017.

[177] Yu-Jin Zhang : An overview of image and video segmentation in the last 40 years. Advances in Image and Video Segmentation, pages 1–15, 2006.

[178] Guoyan Zheng : Statistical shape model-based reconstruction of a scaled, patient-specific surface model of the pelvis from a single standard ap x-ray radiograph. Medical physics, 37(4):1424–1439, 2010.