Fully Automatic CNN-Based Personalized 3D Femur Reconstruction from EOS 2D Bi-Planar Radiographs

by

Nahid BABAZADEH KHAMENEH

THESIS
PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLEMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY IN ENGINEERING
PH.D.

MONTREAL, APRIL 20, 2023

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Mr. Carlos Vazquez, Thesis Supervisor
Department of Software and Information Technology Engineering, École de technologie supérieure

Mr. Jacques A. de Guise, Thesis Co-supervisor
Department of Systems Engineering, École de technologie supérieure

Mr. Ismail Ben Ayed, President of the Board of Examiners
Department of Software and Information Technology Engineering, École de technologie supérieure

Mr. Christian Desrosiers, Member of the jury
Department of Software and Information Technology Engineering, École de technologie supérieure

Mr. Chris Joslin, External Evaluator
Department of Information Technology, Carleton University

THIS THESIS WAS PRENSENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND PUBLIC

APRIL 6, 2023

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

FOREWORD

This doctorate was carried out at the Imaging and Orthopedics Research Laboratory (LIO, ÉTS, CRCHUM). It was co-funded by the the Natural Science and Engineering Research Council of Canada (NSERC), MEDTEQ-MITACS program, and EOS Image Inc., and the École de Technologie Supérieure. Collaboration with orthopedic surgeon, Dr. Frédéric Lavoie (MD, CHUM hospital in Montreal) and his team allowed the retrospective collection of patient data.

ACKNOWLEDGMENTS

To begin with, I would like to express my sincere appreciation to the members of the jury for honoring me by having accepted to evaluate my thesis and attend my defense.

I would like to express my deepest gratitude to my research supervisor, Prof. Carlos Vazquez, for his insightfulness, invaluable advice, enthusiastic encouragement, continuous support, positive attitude, which gave me the courage and confidence to persevere and progress in my research. I am extremely grateful to you for sharing with me your scientific expertise, immense knowledge, and your human qualities. This endeavor would not have been possible without you. I would like to express my sincere appreciation to my co-supervisor, Prof. Jacques A. de Guise, for his thoroughness, resourcefulness, and continuous support, which helped me think in different perspectives and widen my horizons. Accomplishing this thesis would not have been possible without patient guidance, enthusiasm for the project, and continuous support from both of my supervisors.

I wish to sincerely thank Thierry Cresson, for his insightful comments and invaluable feedback in my research. I am thankful to Dr. Frédéric Lavoie for his fruitful collaboration. An extra thanks to Jacques, Thierry, and Carlos for creating such an amazing environment of research and collaboration in Laboratoire de recherche en imagerie et orthopédie (LIO).

I am thankful to all members of LIO, especially Thierry, Lulu, Benjamin, Christine, Tammy, Caroline, and Gerald. Working with such a smart, hard-working, and friendly group of people has been an amazing experience for me.

This dissertation is sincerely dedicated to my mother, Razieh, and my father, Sadegh, who have supported and helped me in all aspects of my life. No words can describe how much I love and respect both of you. I am sincerely thankful to my father and my mother for their precious support and efforts. Specifically, I am extremely grateful to my lovely mother who has always put my preference and comfort before hers. You always were beside of me in my difficult time, and you always supported, encouraged, and inspired me. My dear mother, I would like to express my sincere appreciation to you for your precious presence and your pure and eternal love in my life. I could not have undertaken this journey to realize my big dream without you. My dear mother, I am deeply gratitude to you and I love you so much forever. I am so thankful to my brother, Gholam Reza, for his precious moral support.

A very special thank to the love of my life, Denis, who has been a great companion in this journey. You had a significant contribution to improvement on my French language in Montréal, Québec. Thank you for always being there for me!

Finally, I would like to thank the organizations that funded this PH.D., the Natural Science and Engineering Research Council of Canada (NSERC), MEDTEQ-MITACS program, and EOS Image Inc., the CRCHUM, and the École de Technologie Supérieure.

Reconstruction personnalisée 3D entièrement automatique de la fémur basée sur (CNN) à partir de radiographies bi-planaires EOS

Nahid BABAZADEH KHAMENEH

RÉSUMÉ

Les mesures géométriques cliniques 3D des os des membres inférieurs, tels que le fémur en position debout, sont cruciales dans la planification préopératoire orthopédique et le suivi des patients. En routine clinique, la reconstruction personnalisée du modèle 3D du fémur est un outil utile pour les médecins afin d'analyser une déformation de forme 3D complexe. Ils utilisent un fémur reconstruit en 3D pour quantifier les mesures géométriques cliniques en 3D telles que la taille, les courbures, les orientations, la rotation femoro-tibial et la torsion fémorale. La reconstruction de modèles osseux 3D basée sur des radiographies bi-planaires 2D offre une alternative efficace à la tomodensitométrie (CT) pour la planification chirurgicale orthopédique et le suivi des patients. Les méthodes de reconstruction de modèles osseux 3D basées sur la tomodensitométrie souffrent d'une dose de rayonnement élevée, de coûts d'acquisition et d'une opération avec le patient en position allongée.

Dans ce processus de reconstruction 3D, le recalage 3D/2D est généralement une tâche essentielle pour établir une relation géométrique entre un modèle 3D antérieur connu et les radiographies bi-planaires 2D d'un patient. Ce processus du recalage comprend la pose 3D et l'estimation de la forme 3D des structures osseuses à partir de seulement deux projections 2D. Les méthodes semi-automatiques, telles que celle employée par le système de reconstruction de modèles 3D EOS®, nécessitent l'intervention manuelle d'un opérateur pour l'initialisation de la pose et l'ajustement de la forme et de l'échelle du modèle 3D aux images. Ces interventions manuelles ont un impact sur la précision, la rapidité et la reproductibilité des approches. Dans cette thèse, nous développons une reconstruction fémorale 3D entièrement automatique personnalisée basée sur des radiographies bi-planaires EOS® 2D via des approches basées sur l'apprentissage en profondeur et des réseaux neuronaux a convolution (CNN).

Le flux de travail entièrement automatique et personnalisé de reconstruction du fémur 3D développé traverse deux étapes principales. Premièrement, la pose osseuse 3D et l'estimation de l'échelle isotrope, et deuxièmement, la recalage non rigide 3D/2D. Dans un premier temps, une méthode automatique de recalage de similarité 3D/2D de grossier à fin, est proposée pour recaler automatiquement un modèle 3D générique du fémur dans des radiographies bi-planaires EOS® 2D acquises avec deux champs de vision différents, soit le corps entier et les membres inférieurs entiers ainsi que les orientations des patients en $0^{\circ}/90^{\circ}$ et $45^{\circ}/45^{\circ}$.

Lors d'une première étape, une segmentation sémantique basée sur CNN suivie d'un enregistrement basé sur PCA initialise la pose 3D et l'échelle isotrope du fémur. Ensuite, les régresseurs basés sur CNN affinent les paramètres de pose 3D. La deuxième étape traite de la déformation de la forme 3D locale et de l'échelle 3D en fusionnant le déplacement 3D local basé sur CNN et l'estimation des rapports d'échelle 3D locaux de 17 poignées MLS avec la déformation des moindres carrés mobiles (MLS) pour mieux s'adapter aux radiographies du patient.

Pour valider la première étape, les erreurs de pose 3D et de mise à l'échelle isotrope du fémur sont validées par rapport à des modèles 3D personnalisés de référence floue du fémur reconstruits via un outil logiciel commercial semi-automatique, SterEOS. Dans la deuxième étape, la précision de la forme 3D locale, de l'échelle 3D et du modèle 3D personnalisé de l'ensemble du fémur est validée sur deux ensembles de validation différents.

Le premier ensemble de validation comprend 15 étalons-or flous des modèles 3D personnalisés des fémurs reconstruits via une méthode semi-automatique dans SterEOS. La moyenne et les écarts types (moyenne ± STD) de la racine carrée moyenne des erreurs de distance point à surface (RMS-P2S) sont (0.88±0.29) mm. Le deuxième ensemble de validation comprend 5 modèles 3D reconstruits personnalisés à base de tomodensitométrie de référence du fémur. La (moyenne ± STD) des erreurs RMS-P2S est (2.70±0.39) mm.

Mots-clés: modèle de fémur 3D géométrique personnalisé, radiographies bi-planaires EOS® 2D, estimation de la pose 3D et de l'échelle isotrope, la recalage 3D/2D non rigide, recalage 3D/2D basé sur l'apprentissage en profondeur, régression basée sur le réseaux neuronaux convolution (CNN) modèle, déformation MLS

Fully Automatic CNN-Based Personalized 3D Femur Reconstruction from EOS 2D Bi-Planar Radiographs

Nahid BABAZADEH KHAMENEH

ABSTRACT

Clinical 3D geometric measurements of the lower limb bones such as the femur, in standing position, are crucial in orthopedic pre-operative planning and patient follow-up. In clinical routine, the personalized 3D model reconstruction of the femur is a useful tool for physicians to analyze a complex 3D shape deformation. They use a 3D reconstructed femur to quantify clinical 3D geometric measurements such as size, curvatures, orientations, femoral-tibia rotation, and femoral torsion. 2D bi-planar radiographs-based 3D bone model reconstruction provides an efficient alternative to Computerized Tomography (CT) for orthopedic surgical planning and patient follow-up. CT-scan-based 3D bone model reconstruction methods suffer from high radiation dose, acquisition costs, and operation with patient in reclining position.

In 2D bi-planar radiographs-based 3D model reconstruction, 3D/2D registration is an essential task to establish a geometric relationship between a known prior 3D model and a patient's 2D bi-planar radiograph. This registration process includes the 3D pose and the 3D shape estimation of bone structures from only two 2D projections, is highly complex due to information loss during 2D projection of 3D bone and the need to solve an inverse problem using 2D projected sparse data. Semi-automatic methods, such as the one employed by the EOS® 3D model reconstruction system, require the manual intervention of an operator for the pose initialization and the shape and scale adjustment of the 3D model to the images. These manual interventions impact the accuracy, time-efficiency, and reproducibility of the approaches. In this thesis, we develop a fully automatic EOS® 2D bi-planar radiographs-based personalized 3D femur reconstruction via deep learning approaches.

The developed fully automatic personalized 3D femur reconstruction workflow crosscuts two main stages. Firstly, 3D bone pose and isotropic scale estimation, and secondly, 3D/2D non-rigid registration (3D shape deformation). In the first stage, an automatic coarse-to-fine 3D/2D similarity registration method is proposed to automatically register a generic 3D model of the femur into EOS® 2D bi-planar radiographs acquired with two different fields of view, full body and whole lower limbs, and patients' orientations in $0^{\circ}/90^{\circ}$ and $45^{\circ}/45^{\circ}$. Firstly, a CNN-based semantic segmentation followed by a PCA-based registration is used to initialize the femur's 3D pose and isotropic scale. Then, CNN-based regressors refine the 3D pose parameters. Then, the second stage deals with local 3D shape and 3D scale deformation by merging CNN-based local 3D displacement and local 3D scale ratio estimation of 17 handles with the Moving Least Square (MLS) deformation to get better fit to the patient's radiographs.

To validate the first stage, the 3D pose and isotropic scaling errors of the femur are validated in comparison to fuzzy gold standard personalized 3D models of the femur, reconstructed by an expert via a semi-automatic commercial software tool, SterEOS. In the second stage, the accuracy of local 3D shape, 3D scale, and the personalized 3D model of the whole femur is validated on two different validation sets. The first validation set includes 15 fuzzy gold standards of the personalized 3D models of the femur reconstructed via the semi-automatic SterEOS. The mean and standard deviations (mean±STD) of Root Mean Square of point-to-surface distance errors (RMS-P2S) is (0.88±0.29) mm. The second validation set comprises 5 gold standard personalized CT-scan-based reconstructed 3D model of the femur. The (mean±STD) of RMS-P2S errors is (2.70±0.39) mm.

Keywords: personalized geometric 3D femur model, EOS® 2D bi-planar radiographs, 3D pose and isotropic scale estimation, 3D/2D non-rigid registration, deep learning-based 3D/2D registration, convolutional neural network (CNN)-based regression model, MLS deformation

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF ABREVIATIONS

AI	Artificial Intelligence
CCN	Convolutional Neural Network
2D	Two Dimensional
3D	Three Dimensional
DL	Deep Learning
Fr & LAT	Frontal & Lateral View
ML	Machine Learning
DRR	Digitally Reconstructed Radiograph
PA	Posterior Anterior (2D radiography plane)
PCA	Principal Component Analysis
NSCP	Non Stereo Corresponding Point
NSCC	Non Stereo corresponding Contours
SCP	Stereo Corresponding Point
SSM	Statistical Shape Model

INTRODUCTION

Clinical three-dimensional (3D) geometric measurements of the lower limb bones such as the femur, in standing position, are crucial in orthopedic pre-operative planning and patient follow-up (Cretu et al., 2018 ; van Drongelen et al., 2020). In clinical routine, the personalized 3D model reconstruction of the femur is a useful tool for physicians to analyze a complex 3D shape deformation and quantify clinical 3D geometric measurements such as the size, curvatures, or orientations of the femur (Hosseinian & Arefi, 2015 ; Reyneke et al., 2019).

In personalized 3D femur reconstruction, the main existing approaches are Computerized Tomography (CT)-scan-based (Anderst, Zauel, Bishop, Demps, & Tashman, 2009), Magnetic Resonance Imaging (MRI)-based (Abebe et al., 2011), and 2D bi-planar radiographs-based 3D model reconstruction methods (Baka et al., 2011 ; Youn, Park, & Lee, 2017 ; Yu, Chu, Tannast, & Zheng, 2016). CT-scan and MRI-based methods suffer from high radiation dose and acquisition costs, respectively (Baka et al., 2012). These drawbacks of CT-scan and MRI drive efforts toward reducing 3D acquisitions of CT-scan and MRI. To eliminate radiation dose and acquisition costs, state-of-the-art methods make efforts to reconstruct personalized 3D model of the bone structures from 2D bi-planar radiographs (Baka et al., 2012 ; Yu et al., 2016). In contrast to CT-scan-based 3D reconstruction that operates with the patients in a reclining position (Baka et al., 2012), 2D bi-planar radiographs-based 3D femur reconstruction method provides natural standing position of the patient and a low-level of radiation dose (Chaibi et al., 2012). The 2D bi-planar radiographs-based 3D femur reconstruction provides an efficient alternative to CT (Hosseinian & Arefi, 2015 ; Reyneke et al.1, 2019) for orthopedic surgical planning (Cerveri, Belfatto, & Manzotti, 2020) and patient follow-up (Abebe et al., 2011). In the 3D reconstruction process, 3D/2D registration is essential to establish a geometric relationship between a known prior 3D model and a patient's 2D bi-planar radiographs (Baka et al., 2011 ; Chaibi et al., 2012 ; Cresson, Branchaud, Chav, Godbout, & de Guise, 2010 ; Goswami & Kr., 2015 ; Hosseinian & Arefi, 2015 ; Reyneke et al., 2019 ; Yu et al., 2016). This 3D/2D registration process is highly complex due to information loss during 2D projection of 3D bone and the need to solve an inverse problem using 2D projected sparse data (Baka et al., 2011 ; Chaibi et al., 2012 ; Youn et al., 2017 ; Yu et al., 2016).

Semi-automatic methods, such as the one employed by the EOS® 3D model reconstruction system (Chaibi et al., 2012), require manual interventions, which impact the accuracy, time efficiency, and reproducibility of the 3D femur reconstruction (Chaibi et al., 2012). Recently, state-of-the-art methods were introduced to automate the 3D model reconstruction process (Aubert et al., 2019 ; Yu et al., 2016). However, in 3D femur reconstruction applications, there is still a clinical need for an automatic method which provides physicians, if required, with the means to further manually and easily correct potential errors of the reconstructed 3D femur (Chaibi et al., 2012 ; Yu et al., 2016). To address this problem, in this thesis, we present a fully automatic 3D femur reconstruction method to fit a generic 3D femur model into the patient's EOS® bi-planar radiographs and accurately measure clinical 3D geometrical parameters in a time-efficient manner, while retaining the capacity to quickly and easily adjust the reconstructed 3D femur's possible errors. The EOS® Imaging Inc. is the project partner and the presented project aims to automate the current semi-automatic tool integrated for 3D femur reconstruction.

This thesis is presented in 7 chapters. CHAPTER 1 describes fundamental notions of the context of study. CHAPTER 2 provides a comprehensive literature review on the most relevant state-of-the-art methods to the context of the project. Research problematic and objectives of thesis are driven in CHAPTER 3. The proposed methodology is presented in CHAPTER 4. The experimental setup and validation results are drawn in CHAPTER 5. A comprehensive discussion on each step of the proposed methodology, separately, is illustrated in CHAPTER 6. At the end, CHAPTER 7 outlines conclusions, contributions, and recommendation for the future work.

CHAPITRE 1

FONDAMENTAL NOTIONS

This chapter is organized in six main sections as follows. First, the femur structure is presented. Section 1.2 introduces the femur's shape deformities, followed by section 1.3 presenting orthopedic pre-operative planning. Section 1.4 describes the medical diagnosis and 3D model-based orthopedic applications. Section 1.5 provides a definition on personalized 3D femur reconstruction followed by the associated problem and difficulties. Conclusions on limitation of patient-specific 3D bone are drawn in section 1.6.

1.1 Anatomy of the femur

The human body's lower limb includes three main regions, femur, tibia, and foot. The femur is the strongest bone in the human body. Figure 1.1 illustrates the posterior and anterior of the femur bone structure along with the names of the region on the femur.

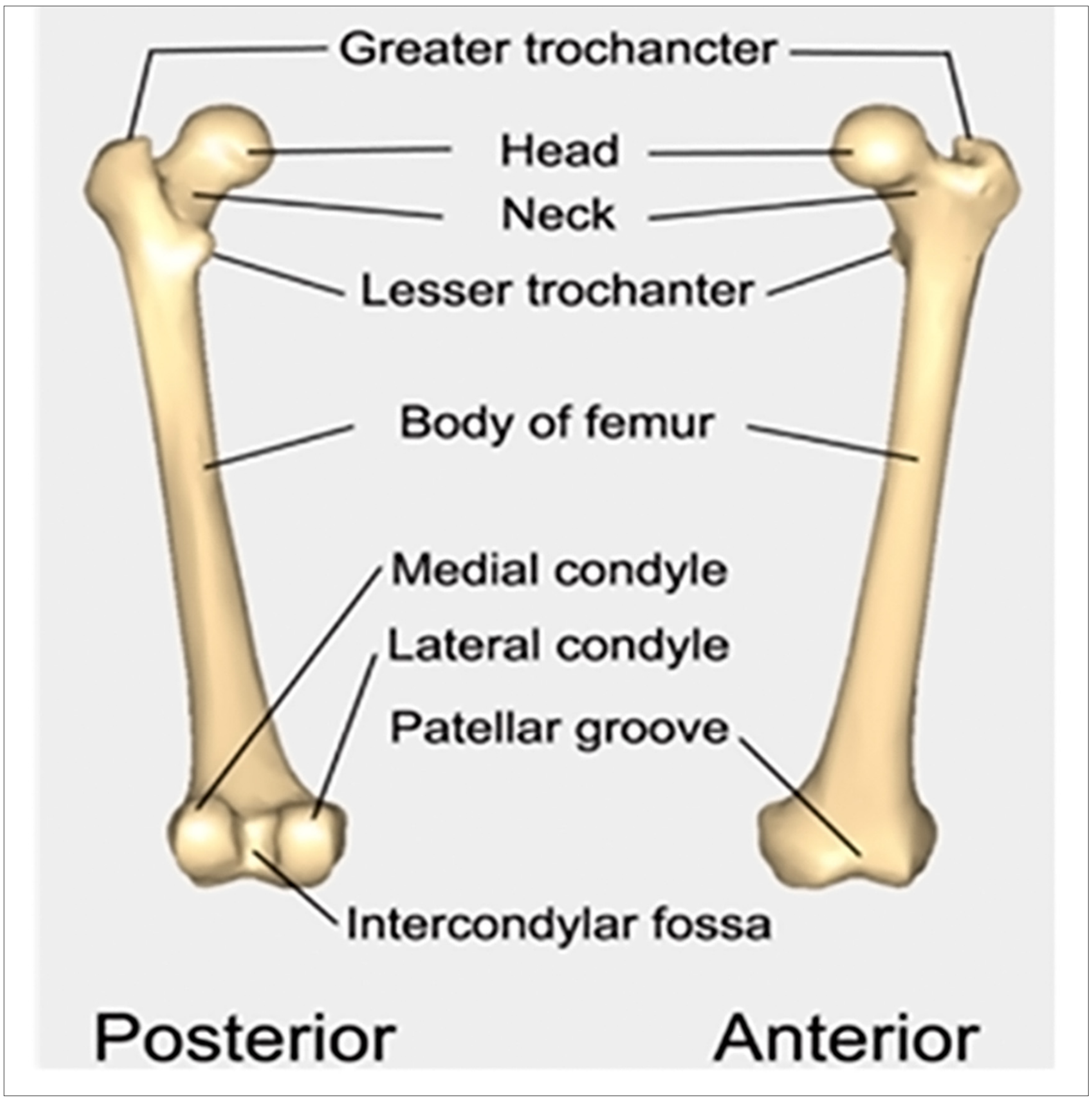


Figure 1.1 Illustration of posterior and anterior of the femur bone structure
Taken from
https://courses.lumenlearning.com/ap1x94x1/chapter/the-lower-limbs

1.2 Femur's shape deformation and clinical 3D geometrical parameters

To analyze a complex 3D shape deformation of the femur, clinical 3D geometric measurements of the femur (Figure 1.2), in standing position, are crucial in orthopedic pre-operative planning and patient follow-up (Cretu et al., 2018 ; van Drongelen et al., 2020). In clinical routine, quantification of the femur's 3D geometric measurements such as size, curvatures, orientations, and femoral torsion (Figure 1.2) provides for physicians essential information (Hosseinian & Arefi, 2015 ; Reyneke et al., 2019).

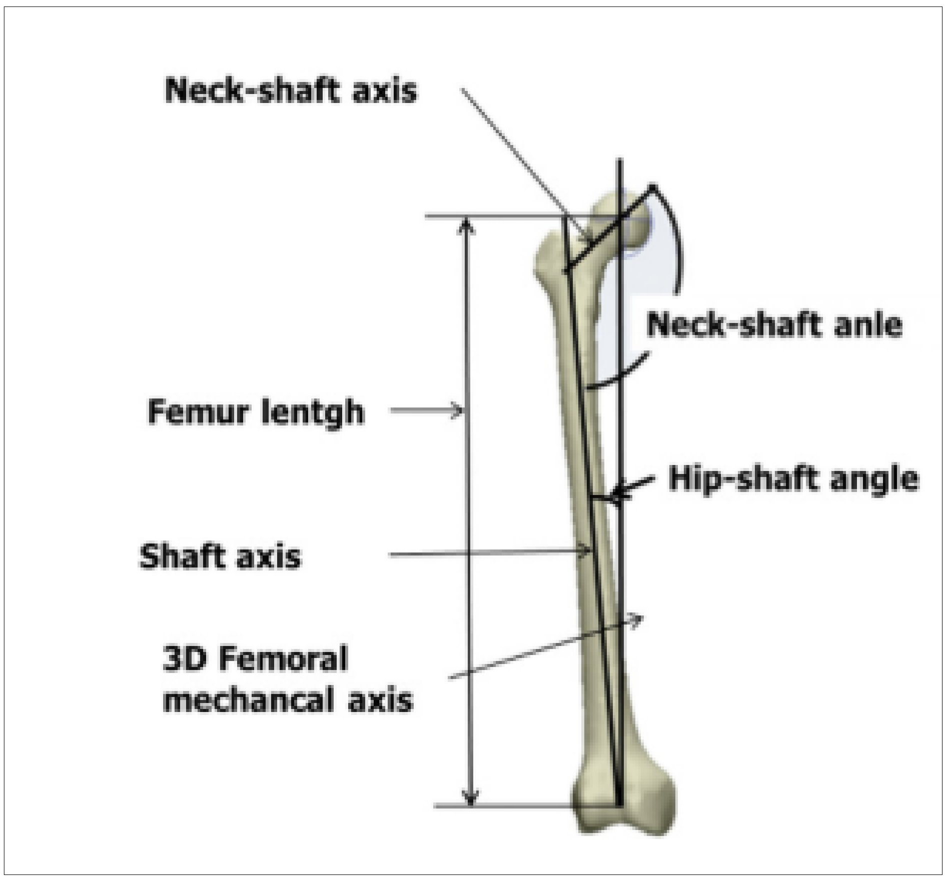


Figure 1.2 Clinical femur 3D geometrical parameters
Taken from van Drongelen et al. (2020)

The most important 3D geometrical parameters corresponding to the femur's shape deformation (Figure 1.3) consist of:

- Hip knee center-femoral shaft angle (HKS), which is the angle between the mechanical and anatomical femoral axis;
- Femoral mechanical angle (FMA),
- Femoral torsion (FT), which is the angle between the femoral neck axis and the bicondylar femoral axis;
- Femoral length (FL), (van Drongelen et al., 2020).

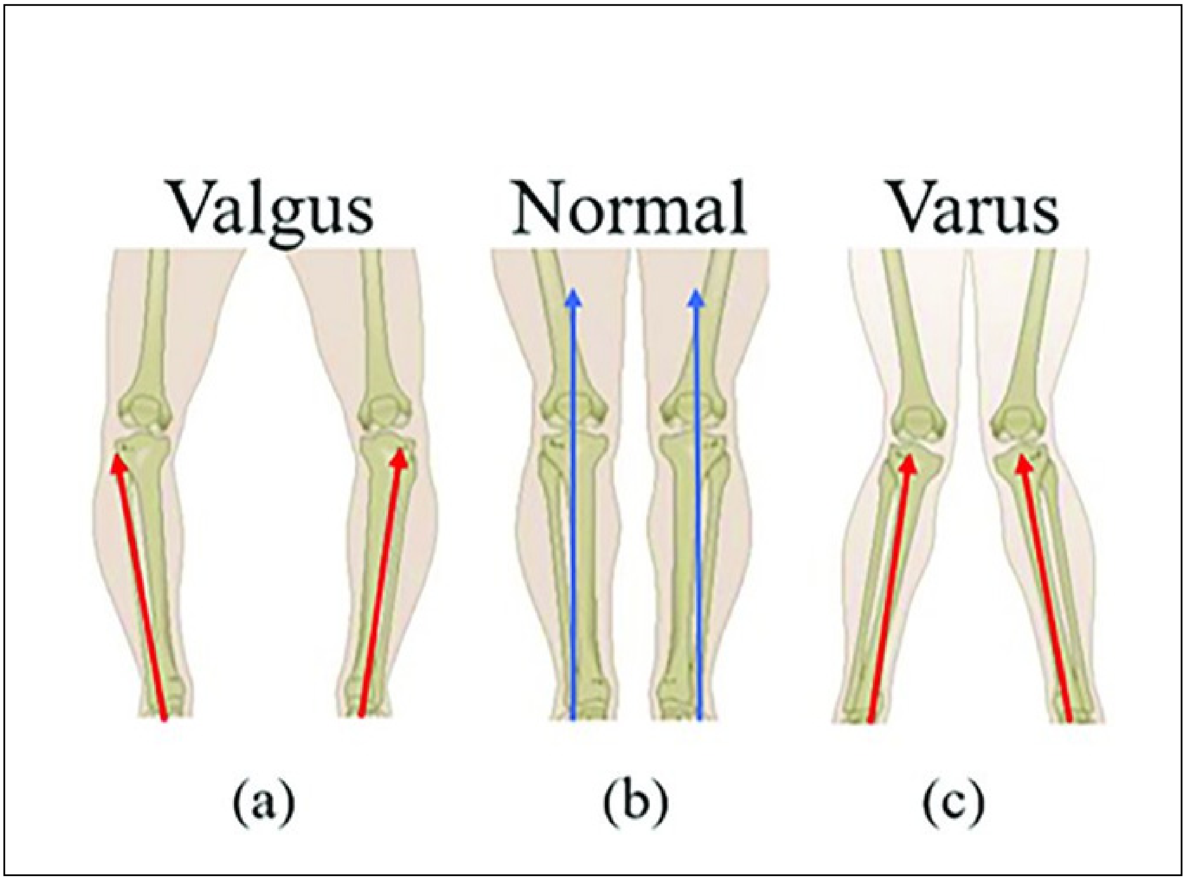


Figure 1.3 Illustration of the femur's shape deformation
Taken from
Rodrigues, Ó Catháin, O'Connor, & Murray (2020)

1.3 Importance of personalized 3D model reconstruction of the femur

In clinical routine, the personalized 3D model reconstruction of the femur is a useful tool for physicians to analyze a complex 3D femur shape deformation. They use a personalized 3D reconstructed model to quantify clinical 3D geometrical parameters of the femur (Hosseinian & Arefi, 2015 ; Reyneke et al., 2019). In comparison of the geometric 3D model of the femur with 2D radiograph, only 3D model provides some important 3D geometric measurements such as femoro-tibial rotation and femoral anteversion (Guenoun, Zadegan, Aim, Hannouche, & Nizard, 2012). Using only a 2D radiograph in which physicians mentally determine the 3D bone deformities without any 3D measurements is not accurate enough in the orthopedic pre-operative planning (Pavan Gamage, Xie, Delmas, & Xu, 2010). Statistical analysis of the clinical geometrical 3D parameters of the femur shows that the geometric 3D model of the femur, in comparison with the 2D radiograph, provides more reliable and reproducible inter- and intra-observer assessments of the femur's clinical 3D parameters such as Femur length, Femoral offset, and (HKS) angle (Guenoun et al., 2012). Table 1.1 compares mean inter-observer differences between two operators for the femur' 3D geometrical parameters measured by the EOS® 2D bi-planar radiograph-based reconstructed 3D femur and the EOS® 2D bi-planar radiographs (Guenoun et al., 2012). Table 1.1 shows that using the 3D femur model rather than 2D bi-planar radiographs lead to decrease mean of the inter-observer differences of the 3D geometric parameters measurements (Guenoun et al., 2012).

Table 1.1 Comparison of statistical results of 3D geometrical Parameters of the femur
Adapted from Guenoun et al. (2012)

	Mean Inter-observer differences via EOS® 3D femur reconstruction	Mean Inter-observer differences via EOS® 2D bi-planar radiographs
Femur length (cm)	0.132	0.206
Femoral offset (cm)	0.269	0.312
Femoral head diameter(cm)	0.252	0.359
Femoral neck length (cm)	0.265	0.466
HKS angle (°)	0.519	0.868
HKA angle (°)	0.497	0.519
Neck shaft angle (°)	2.937	4.685

1.4 EOS® imaging system

In 1992, Professor Georges Charpak invented a new gaseous particle X-ray detector and received the Nobel Prize (Melhem, Assi, El Rachkidi, & Ghanem, 2016). A team of orthopedic surgeons, radiologists and biomedical engineers transformed this invention into the new low-dose system called EOS® (EOS® imaging, Paris, France), (Guenoun et al., 2012). The EOS® sensors reduce the level of the radiation dose, 50-80 percent less than conventional radiographs (Melhem et al., 2016). The EOS® cabin system includes a C-shaped arm that vertically travels and supports two orthogonally placed image acquisition systems. Figure 1.4 shows the principle of the EOS® system. The source and detector move together and the beam are horizontal to the patient (Guenoun et al., 2012).

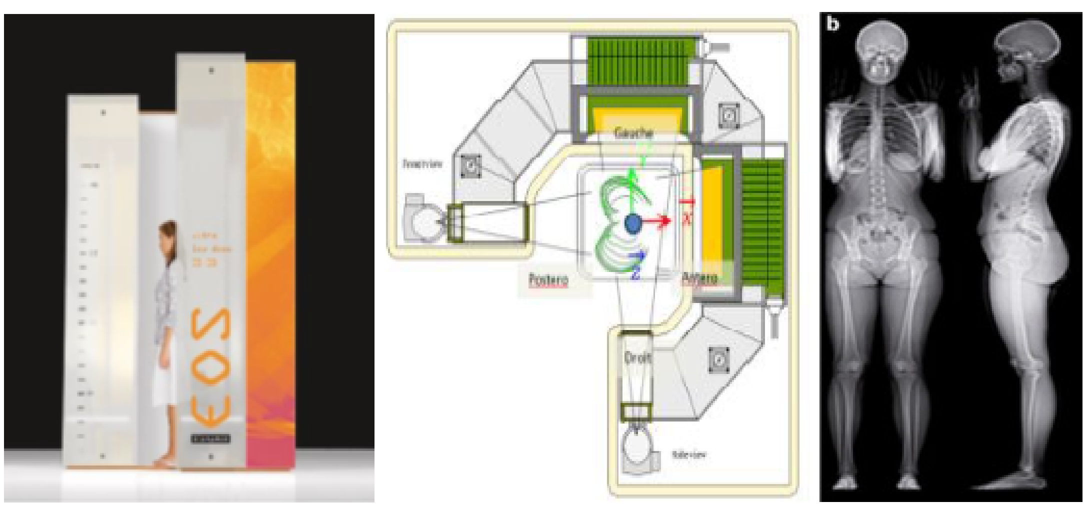


Figure 1.4 Principal of the EOS® cabin system
Modified picture from
https://www.eos-imaging.com/our-expertise/imaging-solutions/eos-system

The EOS® imaging system produces long-length, weight-bearing 2D bi-planar radiographs with minimum irradiation dose (Melhem et al., 2016), Figure 1.5.

Figure 1.5 Illustration of the stereo radiography in EOS® cabin system
Modified picture from
https://universityorthopedics.com/EOS/index.html

1.5 Stereo radiography

Stereo-radiography is an imaging modality consisting of taking x-ray images of a patient from different angles, typically two perpendicular view points (2D bi-planar radiographs), frontal (AP) and lateral (LA), (Figure 1.6). Both frontal and lateral views provide complementary information but the information remains partial because the x-ray images represent the projection of all the organs of the body. In addition, the distance and angle measurements present inaccuracies due to the radiological projection bias (non-constant scales). However, stereovision algorithms can be used to perform 3D modeling if the images are geometrically calibrated.

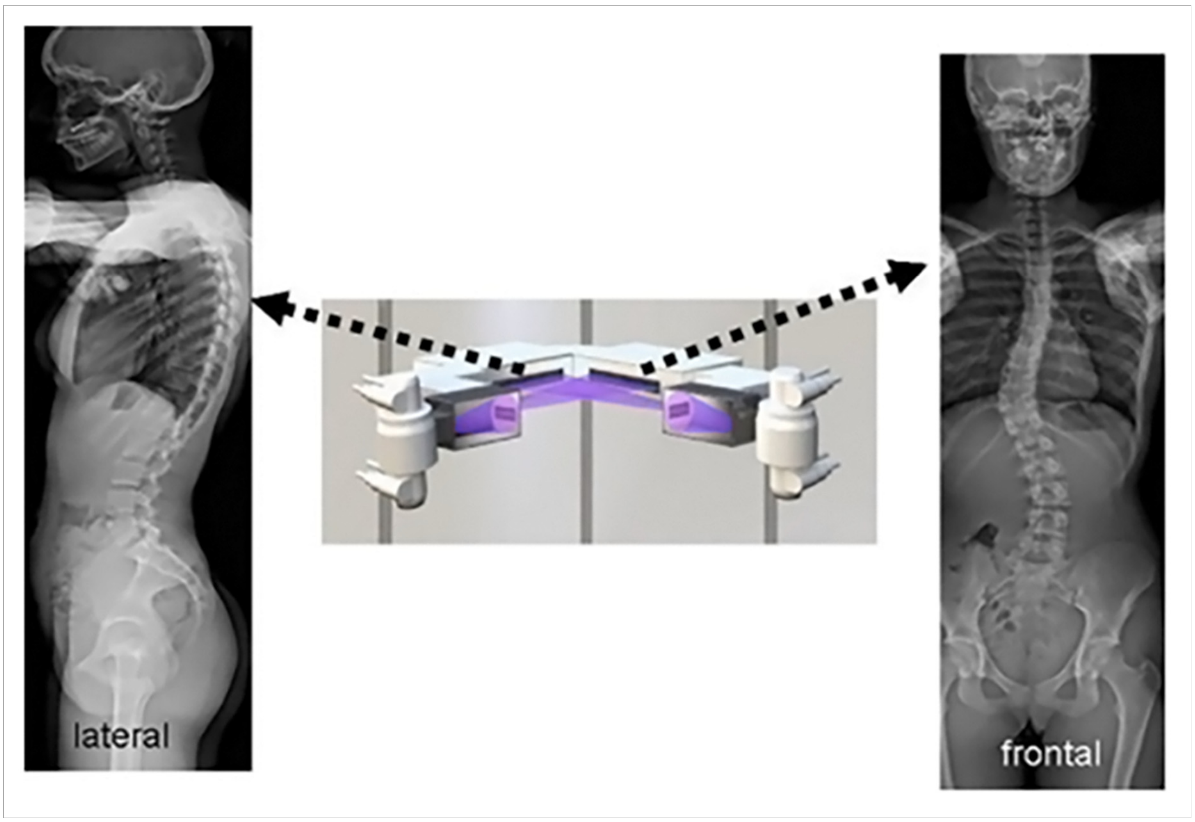


Figure 1.6 Illustration of simultaneous capturing of spatially calibrated frontal and lateral radiographs in EOS® stereo radiography system
Taken from Wybier & Bossard (2013)

1.6 EOS® 2D radiograph-based 3D femur reconstruction

In 3D femur reconstruction, the first prerequisite is to use calibrated bi-planar radiographs with the knowledge of the 3D environment such as: source-detector distance, detector width, collimation, etc. The projection parameters and matrices for projecting a 3D point onto an image or calculating a projection line from a point in the image are assumed to be known. Triangulation methods derived from epi-polar geometry (Mitton et al., 2000) can therefore be used, similar to binocular stereo vision systems, to find the 3D depth of points (Figure 1.7).

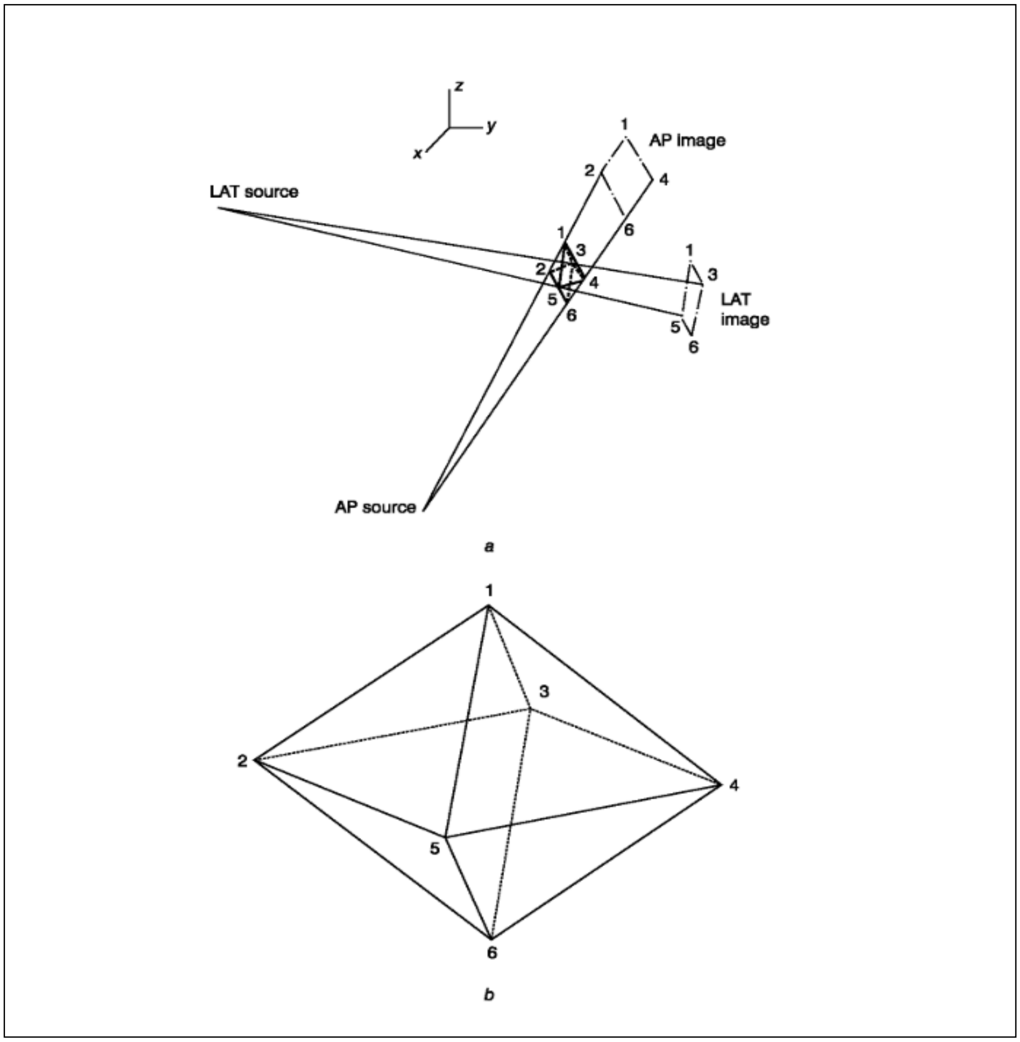


Figure 1.7 (a) Diagram of the stereo radiography to generate (AP) and (LAT) images of a diamond with depth, (b) Close-up view of the diamond
Taken from Mitton et al. (2000)

The EOS® 2D bi-planar radiographs with the patient in standing position (Figure 1.8) present a useful advantage in the investigation of torsion troubles in bony structures, such as the femur, and rotational troubles in articulations with low levels of radiation dose (Chaibi et al., 2012). In the 3D model-based orthopedic applications, the standing position is important to examine the femur deformities with normal weight-bearing situations in order to make standard diagnosis (Cresson et al., 2010).

(A) (B)

Figure 1.8 (A) frontal and (B) lateral EOS® 2D radiograph

As another advantage, the EOS® system provides the possibility of personalized 3D femur reconstruction (Figure 1.9) based on 2D bi-planar radiographs (frontal and lateral views) (Melhem et al., 2016).

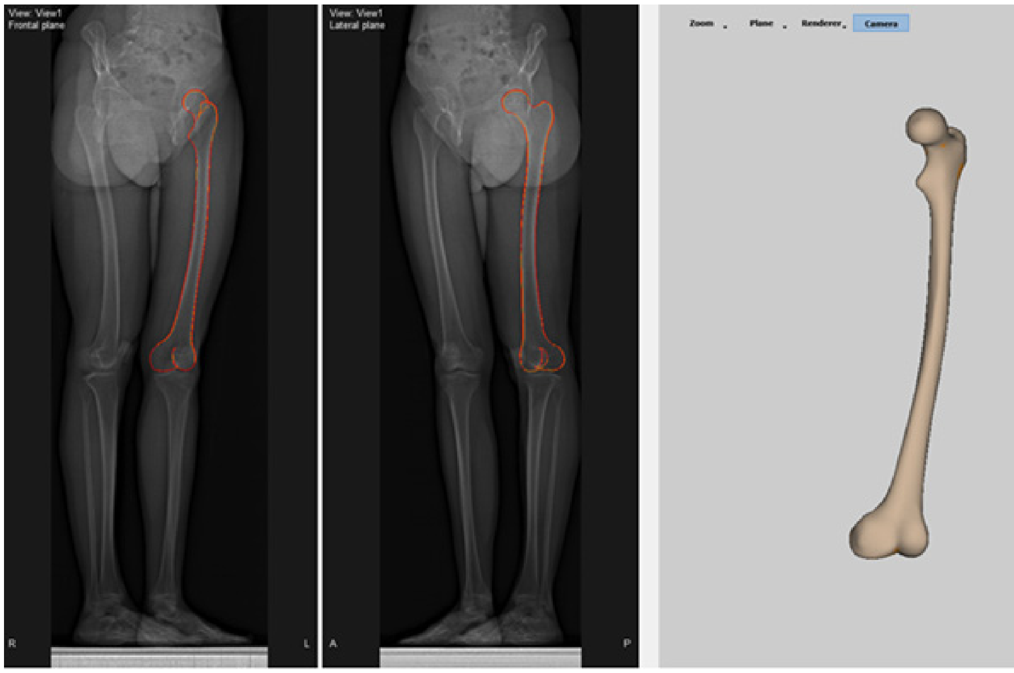


Figure 1.9 EOS® 2D bi-planar radiograph-based 3D femur reconstruction

1.7 Difficulties on EOS® 2D radiograph-based 3D femur reconstruction

However, the 3D femur reconstruction from only two 2D projections is highly complex due to information loss during 2D projection of 3D bone and to the need to solve an inverse problem using 2D projected sparse data. In addition, the difficulties in visualization of the certain parts of the femur in 2D radiographs make the 3D femur reconstruction as a non-trivial task. Some important visualization problems of the femur bone structure in 2D radiographs include:

- Overlapping of the head of the femur of the left and the right lower limbs in the lateral view (Figure 1.10);

- Overlapping of the condyles of the femur of the left and the right lower limbs in both frontal and lateral view (Figure 1.11);
- Difficulty in discrimination between the left and the right femur in the lower limbs in the lateral view (Figure 1.12).

Due to these difficulties in visualization of the femur bone structure in 2D radiographs, the automating the 3D femur reconstruction is very complex and the current 3D femur reconstruction commercial tool (SterEOS 3D) still remains semi-automatic (Chaibi et al., 2012).

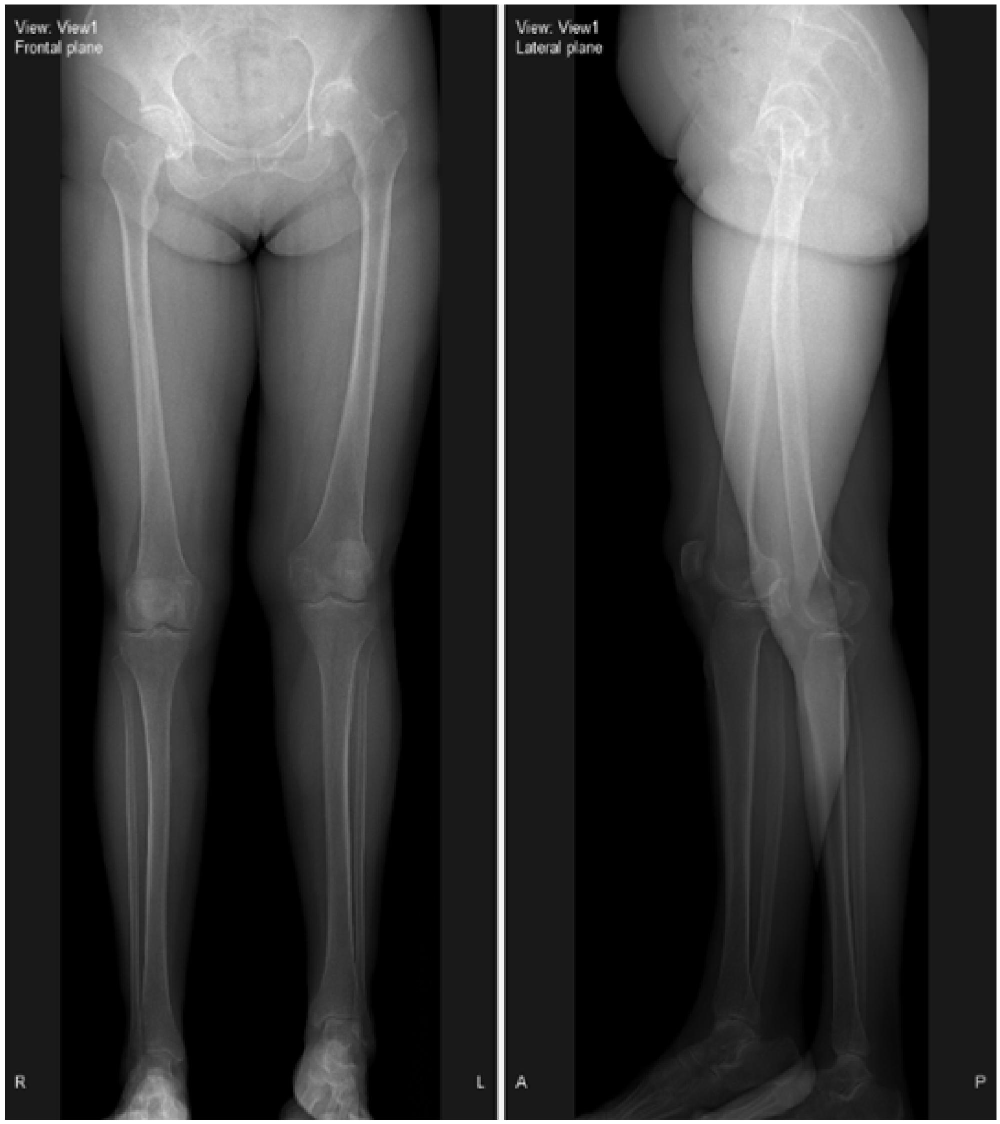


Figure 1.10 Illustration of overlapping of the head of the femur of the left and right lower limbs in lateral view

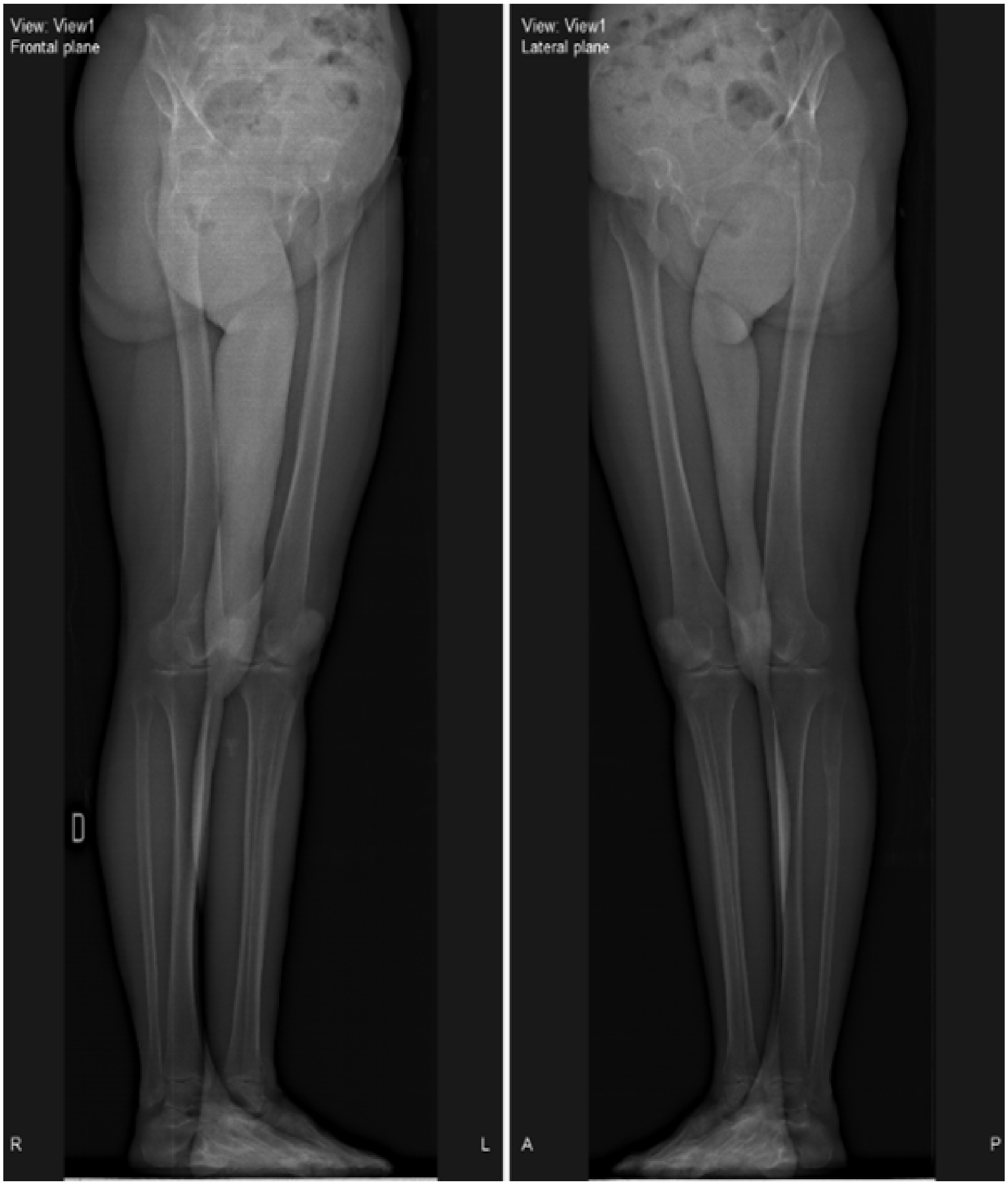


Figure 1.11 Illustration of overlapping of the condyles of the femur of left and right lower limbs in both frontal and lateral view

Figure 1.12 Illustration of difficulty in discrimination between left and right femur in lower limbs in lateral view

1.8 SterEOS 3D

The SterEOS 3D is a commercial software tool, which is deployed in the EOS® 2D bi-planar radiograph-based 3D femur reconstruction platform, (Chaibi et al., 2012 ; Melhem et al., 2016). Figure 1.13 illustrates a schematic of the EOS® 2D bi-planar radiographs-based personalized 3D femur reconstruction software tool, in the laboratory LIO. The semi-automatic SterEOS 3D is currently used in clinical routine, and uses Moving Least Square (MLS) technique, which is a user-friendly 3D femur reconstruction method (Chaibi et al., 2012).

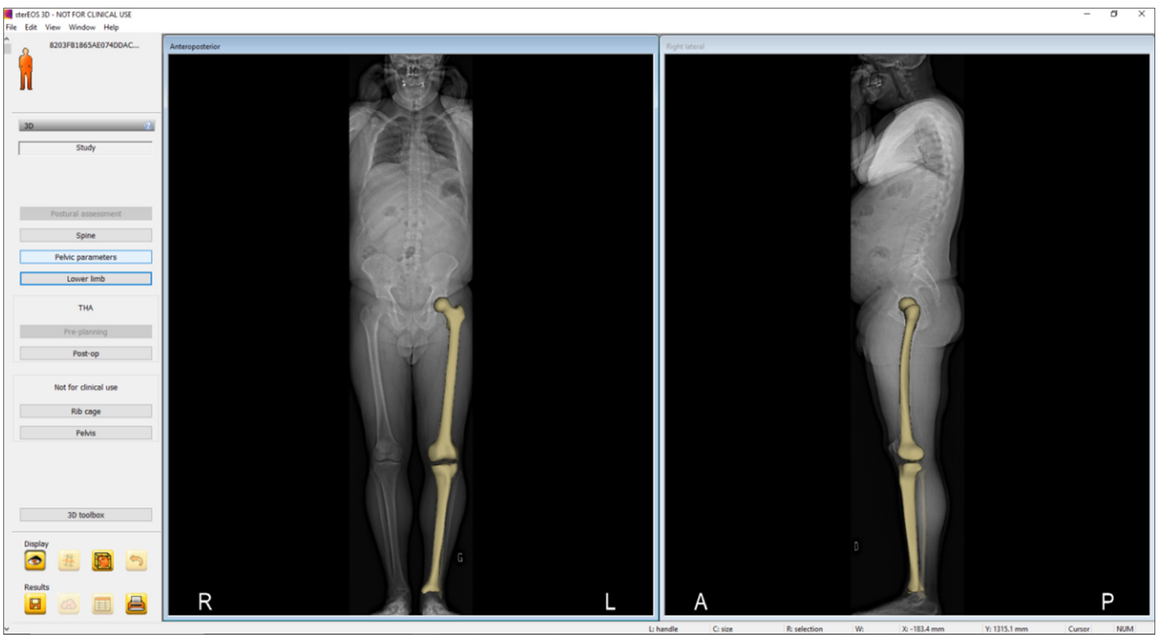


Figure 1.13 Schematic of EOS® 2D bi-planar radiographs-based personalized 3D femur reconstruction software tool

In the SterEOS 3D tool, 3D femur reconstruction process is a tedious work and asks an operator to interact through the 3D femur reconstruction steps:

1) Manual identification of the femur bone structure to initialize the generic 3D model of the femur in the patient's bi-planar radiographs via a manual localization of three MLS handles:

a) Center of the femoral head;

b) Center of the femoral trochlee;

c) 1/3 of diaphysis on the distal femur (Figure 1.14).

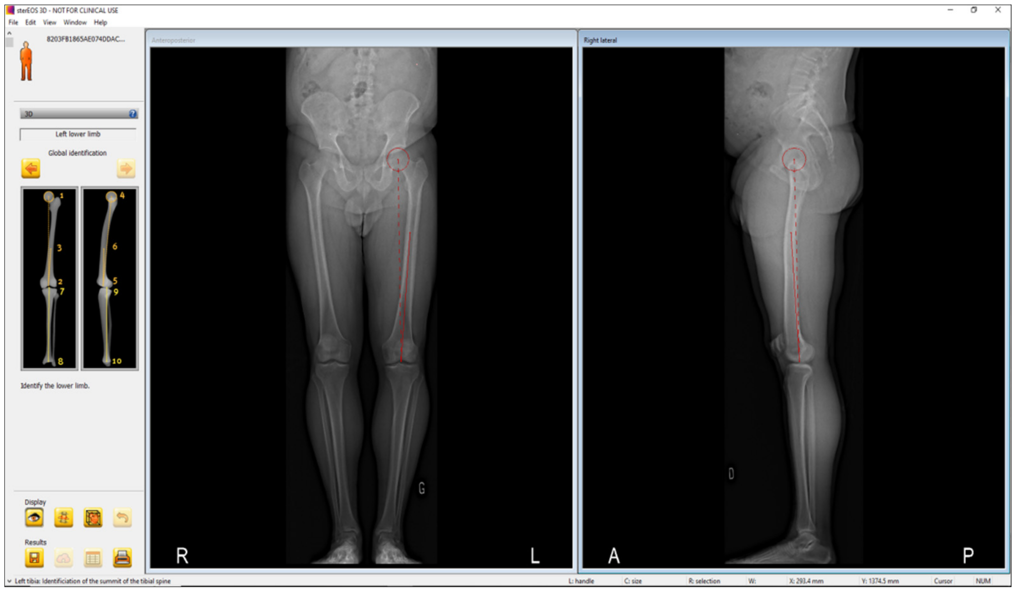


Figure 1.14 Illustration of manual initialization of MLS handles of femur

2) Manual adjustments to refine the positions of the three MLS according to the size and shape of the femur bone structure appeared in patient's 2D radiographs:

a) Adjustment of the center and size of the circle around the MLS handle corresponding to the center of the head of the femur in both frontal and lateral views of patient's radiographs (Figure 1.15);

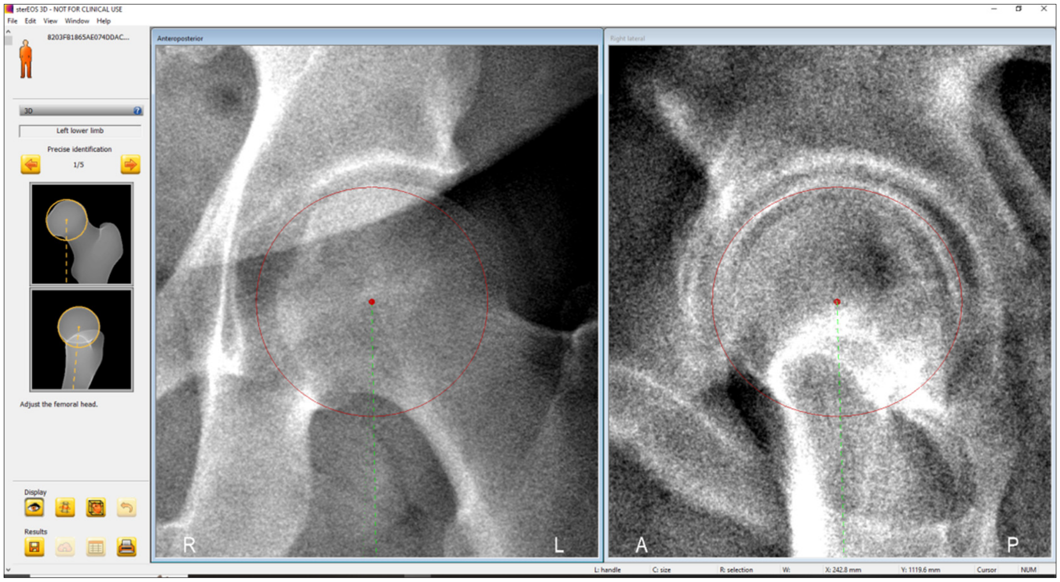


Figure 1.15 Manual adjustment of the position of the center of the femoral head

b) Adjustment of the position of the center of the diaphysis according to both frontal and lateral views of patient’s radiographs (Figure 1.16);

Figure 1.16 Manual adjustment of the position of the center of the diaphysis

c) Manual adjustment of the position of the trochlee, the center and the size of two circles corresponding to the condyle post-interior and the condyle post-exterior according to both frontal and lateral views of patient's radiographs (Figure 1.17).

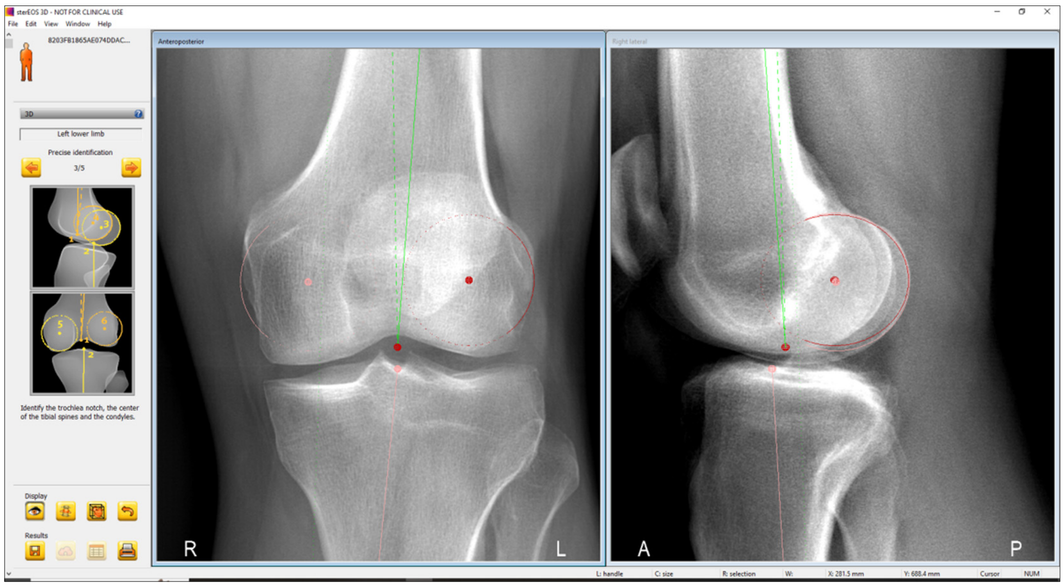


Figure 1.17 Manual adjustment of position of trochlee and size of two circles corresponding to condyle post-interior and condyle post-exterior

3) Adjustment of the 2D projected contours of the initialized 3D model of the femur superimposed on both frontal and lateral views of patient's radiographs (Figure 1.18). An operator manually adjust 17 MLS handles (using the mouse) to deform the shape of the 3D model to get better fit;

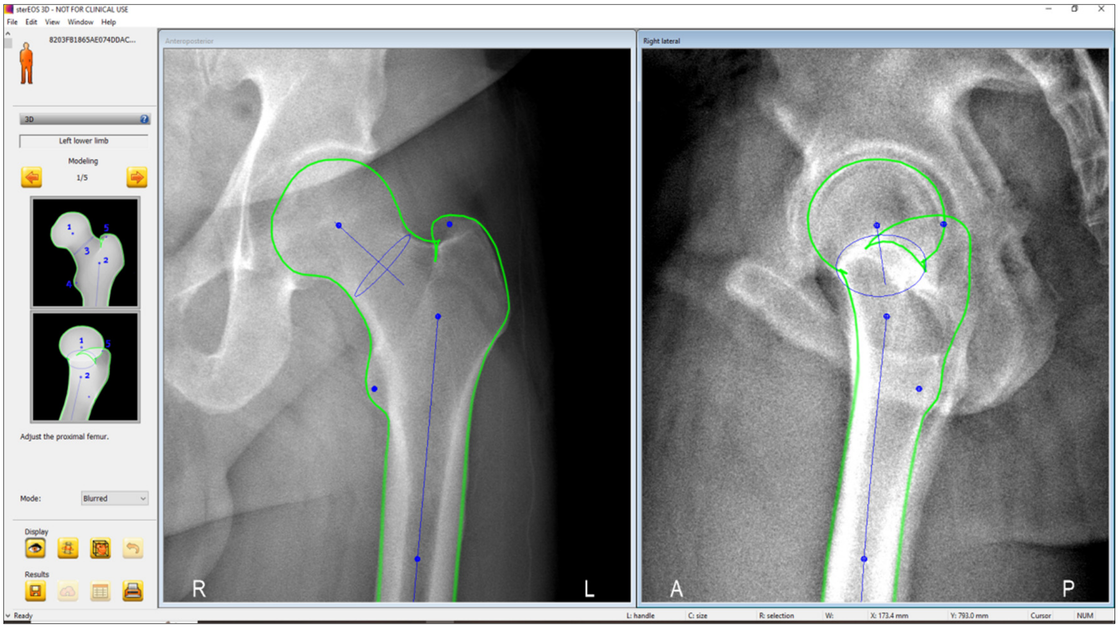


Figure 1.18 Manual adjustment of MLS handles and projected contours of 3D femur

4) Obtaining the final reconstructed 3D model of the femur (Figure 1.19).

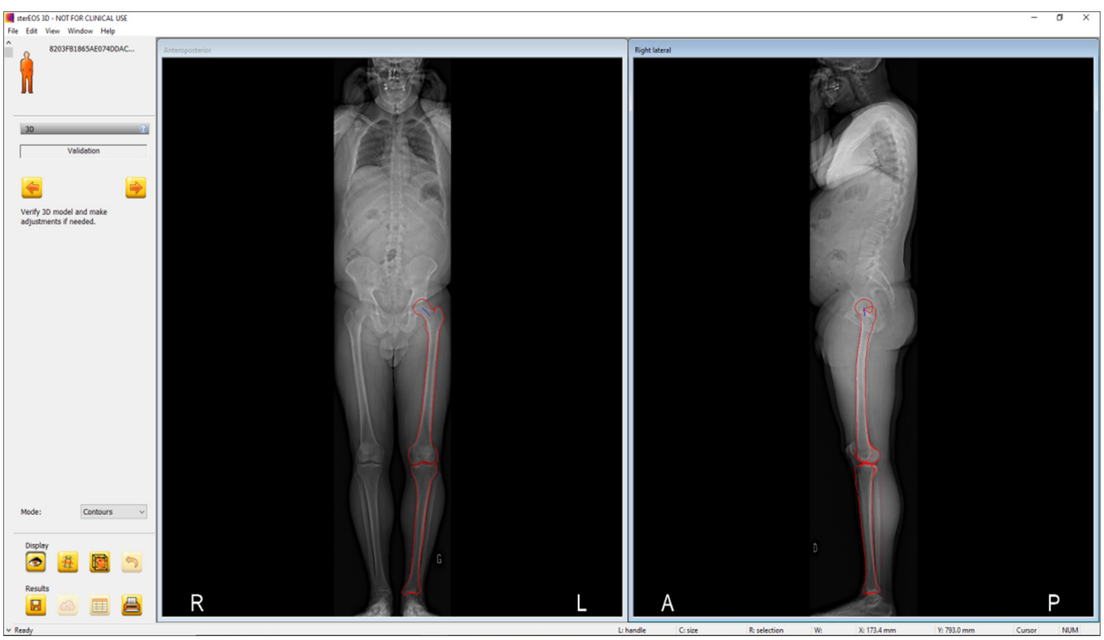


Figure 1.19 Personalized 3D femur superimposed (red color) on patient's radiographs in both frontal and lateral views

Although this semi-automatic approach provides physicians with an easy and user-friendly way to manually carry out adjustments to correct potential 3D reconstruction errors, it has nonetheless certain limitations, such as manual initialization (Figure 1.14), manual adjustment (Figures 1.15, 1.16, 1.17, and 1.18), dependency on the operator's skill, limited reproducibility for lower limbs angular measurements with (95% confidence intervals (CI) $\leq$ 2.0°) and for the neck shaft angle with (95% CI = 4.7°), and high time consumption (10 minutes for both lower limbs reconstruction).

1.9 Conclusion

In summary, 3D model reconstruction of the femur allows carrying out, firstly, innovative applications to analyze and better understand femoral shape deformation, for pre-operative surgical planning and personalized treatments, and secondly quantification of the 3D geometric parameters of the femur for the diagnosis and patient's follow-up. The EOS® 3D model reconstruction system provides an ideal solution for deploying the stereo radiography process in hospitals. The two perpendicular and calibrated images allow 3D femur reconstruction in standing position with a minimal dose of irradiation. In this chapter, we have introduced fundamental notions such as:

- Description of clinical parameters extracted from 3D femur reconstructed models, which are useful to clinicians;
- Stereo radiography and the EOS® imaging system;
- Current semi-automatic EOS® bi-planar radiographs 3D femur reconstruction approach integrated in the commercial SterEOS 3D software tool;
- Advantages and limitations of the current SterEOS 3D tool.

The current semi-automatic SterEOS tool can be automated in order to accurately assess clinical 3D geometric parameters in a time-efficient manner, while retaining the capacity to quickly and easily adjust the reconstructed 3D femur. In the next chapter, we will review the different 3D femur reconstruction algorithms already proposed, as well as artificial intelligence and deep learning based methods.

CHAPITRE 2

LITERATURE REVIEW

2.1 Introduction

3D bone reconstruction methods from 2D radiographs are widely reviewed in (Goswami & Kr., 2015 ; Hosseinian & Arefi, 2015 ; Reyneke et al., 2019). To provide a comprehensive literature review on the state-of-the-art 2D radiograph-based lower limb 3D bone reconstruction methods, this chapter is organized in five main sections. Section 2.2 explains the general workflow of the 3D/2D registration in 3D model reconstruction process used by the state-of-the-art methods. Section 2.3 focuses on 3D pose estimation step describing the evolution of applied methods for manual, semi-automatic, and automatic identification and detection of the 3D bone structures in patient's 2D radiograph. Then 3D/2D rigid or similarity registration approaches are described based on the kind of extracted information from 2D radiographs including geometric, iconic, hybrid, and also deep-learning-based registration methods. Section 2.4 covers state-of-the-art 3D/2D non-registration methods in six sub-sections. Section 2.5 completes the literature review chapter by presenting evaluation protocols comprising evaluation data and evaluation metrics.

2.2 2D radiographs-based 3D bone reconstruction

In orthopedic applications, the state-of-the-art methods address 3D model reconstruction of the bone structures from 2D radiographs via 3D/2D registration in two main stages (Baka et al., 2011 ; Goswami & Kr., 2015 ; Hosseinian & Arefi, 2015 ; Reyneke et al., 2019 ; Yu et al., 2016):

1) 3D pose estimation (3D/2D rigid or similarity registration);
2) 3D/2D non-rigid registration (local 3D shape deformation).

Figure 2.1 illustrates two main stages of the 2D radiographs-based 3D bone reconstruction process. The input of the 3D model reconstruction workflow includes 2D radiographs and a prior 3D model. The output is a personalized 3D model. The 3D reconstruction process includes two main stages: pose estimation and shape deformation. After shape deformation, the 3D reconstruction criteria is evaluated based on a predefined threshold. If the criteria are converged then the 3D reconstruction is done. Otherwise, the shape deformation stage is repeated until the converge criteria are met.

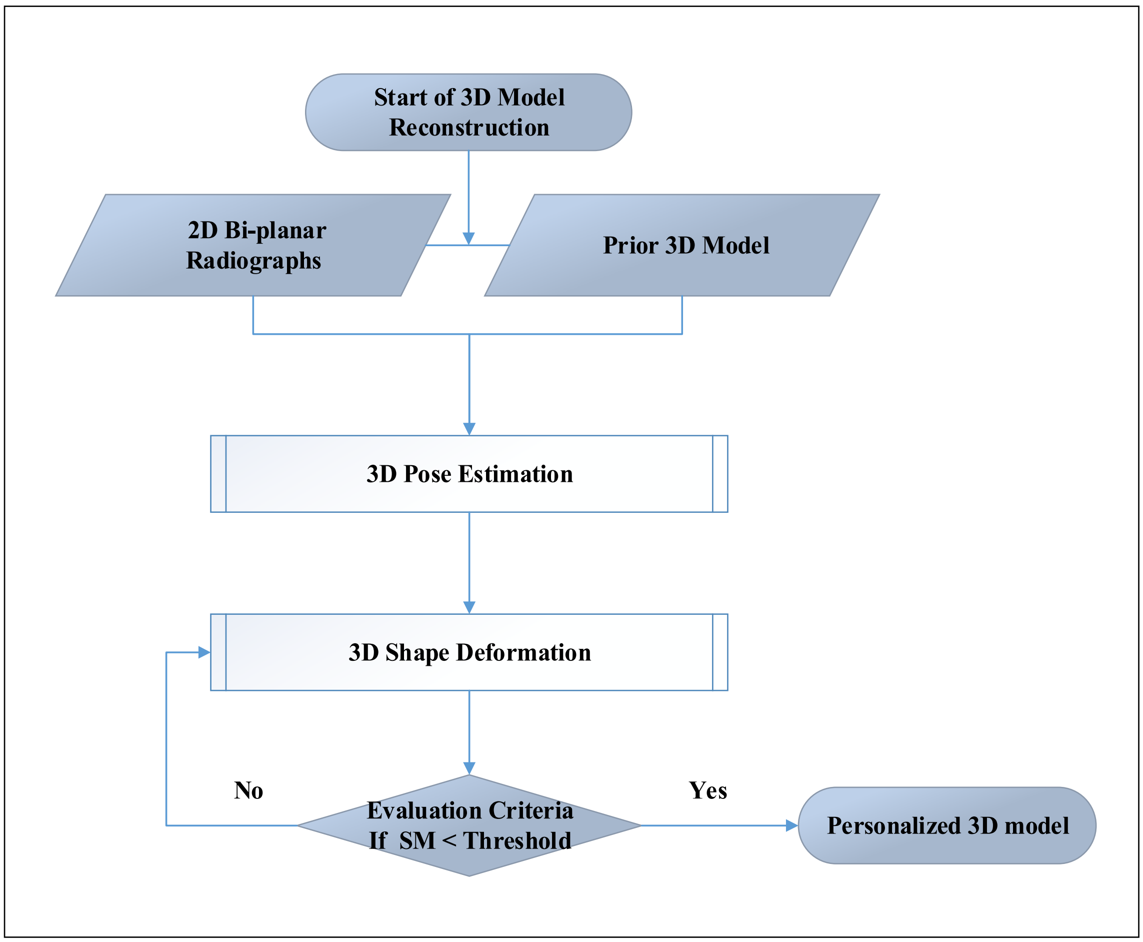


Figure 2.1 2D radiographs-based 3D model reconstruction workflow

2.2.1 3D/2D registration in mathematic formula

The goal of 3D/2D registration is to find a spatial transformation Uthat optimally fits moving or source model (M) into reference or fixed images (F), (Lisa Y. W. Tang and Ghassan Hamarneh, 2013). The 3D/2D registration problem can be mathematically defined as an objective function or energy function comprising two principal terms, matching term and regularization term (Lisa Y. W. Tang and Ghassan Hamarneh, 2013):

$$E(U) = \alpha D(F, M, U) + \sum_{i=1}^{n} \beta_i R_i(U) \tag{2.1}$$

where $E(U)$ is the energy function, also called the objective function of registration problem, D is the data (dis)similarity term (matching term) and R_i is the regularization term, α depicts related weight on the data (dis)similarity term, β_i depicts related weight on the i^{th} regularization term (Lisa Y. W. Tang and Ghassan Hamarneh, 2013). Both- regularization term and data (dis)similarity measurement term play major roles to find optimal transformation U. Data (dis)similarity term measures how well transformation U maps reference and source based on the type of the extracted image data in the matching step. Regularization term is defined based on the displacement fields of local 3D shape deformation. Regularization term measures the regularity of U and provides more constraints on the behavior of U to achieve transformation that is more plausible to obtain local 3D shape deformation. The optimal U can be obtained by minimization of energy function, $E(U)$ combining weighted sum of two or more criteria (Lisa Y. W. Tang and Ghassan Hamarneh, 2013). In general, in Eq. (2.1), the solution of the 3D/2D registration, $U,$ is not unique and registration problem is an ill-posed one (Lisa Y. W. Tang and Ghassan Hamarneh, 2013). State-of-the-art methods make efforts to find optimal transformation parameters as Eq. (2.2), to fit 3D model to reference images.

$$\widehat{U} = \arg min_U \, E(U) \tag{2.2}$$

In a 3D/2D registration, the spatial transformation U combines both global transformation (linear transformation) for 3D pose and scale estimation, and local transformation (non-linear transformation) for 3D/2D non-rigid registration (Cuingnet et al., 2012), so it is formulated as Eq. (2.3):

$$U = L(G_{3D}(V)) \tag{2.3}$$

Where, V is a array of 3D vertices and represents the 3D position of each point of the source 3D model in the world 3D coordinate system as (x, y, z). G_{3D} is linear global transformation (3D pose and scale estimation), and L is nonlinear local transformation (3D/2D non-rigid registration). At first stage, in the global transformation, G_{3D}, we estimate 3D translations, 3D rotations, and isotropic scaling. Then, in second stage for local transformation, L, we estimate and iteratively optimize local 3D shape deformation until convergence.

2.3 3D pose estimation

The 3D Pose estimation process aims to find optimal 3D pose parameters, G_{3D}, i.e., 3D translations, 3D rotations, and isotropic scaling in order to globally fit moving or source 3D model (M) to reference images (F). In the 3D pose estimation stage, we optimize the spatial transformation, called U_{Pose}, by minimizing data (dis)similarity term of the objective function of 3D/2D registration problem, as Eq. (2.4):

$$D(U_{Pose}) = D(F, M\ , U_{Pose}) = D\ (F - U_{Pose}(M)) = D(G_{3D}) \tag{2.4}$$

In the rest of the context, global transformation, linear transformation, and rigid or similarity registration all refer to the 3D pose estimation.

In orthopedic applications, 3D pose estimation methods are widely reviewed in (Markelj, Tomaževič, Likar, & Pernuš, 2012 ; Reyneke et al., 2019). In 2D radiograph-based 3D femur reconstruction application, state-of-the-art 3D/2D registration methods range from semi-automatic (Baka et al., 2011 ; Chaibi et al., 2012 ; Youn et al., 2017) to automatic (Yu et al., 2016) methods. In 3D pose estimation process, following a semi-automatic or automatic identification and detection of the target bone structure in 2D radiograph, a 3D prior model is transformed in 2D reference space, via 3D/2D rigid or similarity registration approaches. This section reviews the target bone identification and 3D/2D rigid or similarity registration methods through following sub sections. Sections 2.3.1, 2.3.2, and 2.3.3 present semi-automatic, automatic identification of the target bone structure and 3D pose estimation initialization methods, and conclusions, respectively. Section 2.3.4, 2.3.5, 2.3.6, and 2.3.7 review geometric feature, intensity, hybrid, and deep learning-based 3D/2D rigid or similarity registration, respectively. Section 2.3.8 outlines conclusions on 3D pose estimation.

2.3.1 Semi-automatic identification of target bone and 3D pose initialization

In semi-automatic 3D pose estimation, a user manually initializes the 3D pose of the prior 3D model either by manually locating the prior 3D model of the bone structures in the 2D radiographs' projection space or manual annotations on the target 3D bone in 2D radiographs. Then, a 3D/2D rigid or similarity registration is used to transform the prior 3D model into 2D radiographs. In (Mitton et al., 2000), a user- manually annotates the non-stereo corresponding points (NSCP), which are visible in only one of the radiographic planes, frontal or lateral 2D radiograph, and then using the 3D position of the non-stereo corresponding points, a generic 3D model is transformed and registered into the 2D bi-planar radiographs. In (Baka et al., 2011, 2012), a user manually initializes the 3D pose of a mean shape of the distal femur close to the optimal position of the target bone in 2D bi-planar radiographs. Then, the user could adjust the 3D model by visualizing the projected silhouette of the distal femur on 2D radiographs. In (Youn et al., 2017), a user roughly annotates 6 points as landmarks on the boundary of the femur in 2D radiographs of the patients. The initial 3D pose of the mean shape of the femur is estimated by finding corresponding 3D points of the six annotated landmarks on frontal and lateral projections. In (Chaibi et al., 2012), to identify the femur in 2D bi-planar radiographs, a user manually locates and adjusts three spheres, as geometric primitive, corresponding to femoral head and two femoral posterior condyles, to represent the main features and global geometry of the femur bone structure. After a rough identification of the target femur, ten non-stereo corresponding points, as anatomical landmarks, are semi-automatically identified and then manually adjusted on the patient's bi-planar radiographs to get better fit and calculate the femoral neck shaft angle (FNSA). Manual identification of the target bone structure in 2D radiographs is time consuming and the accuracy of the annotations and the estimated 3D pose depends to the operator skills. Hence, the semi-automatic 3D pose estimation suffers from the lack of reproducibility. For instance, in (Chaibi et al., 2012), both lower limbs reconstruction takes 10 minutes. The reproducibility for lower limbs angular measurements is (95% confidence intervals (CI) $\leq 2.0^{\circ}$) and for the neck shaft angle with (95% CI = 4.7°). Regarding the shape accuracy, the reported mean and 2 RMS errors are 1.0 and 2.4 mm, respectively, for femur reconstruction.

In another semi-automatic 3D pose estimation method, user interactions are required to identify 3D bone structures in the 2D radiographs' projection space, and then, 3D/2D rigid or similarity registration is used to transform the prior 3D model into 2D radiographs. For instance, in (Laporte, Skalli, de Guise, Lavaste, & Mitton, 2003), for 3D pose initialization of a generic 3D model, 2D non-stereo corresponding contours (NSCC) of the distal femur are extracted from 2D bi-planar radiographs via a semi-automatic active contours-based method. Then, the initial 3D pose of the distal femur is estimated by calculating the corresponding 3D contours, and then computing transformation parameters between the generic model and the corresponding 3D contours to register the generic 3D model into the 2D bi-planar radiographs. Manual intervention of an operator for 3D pose initialization negatively impacts time efficiency. In (Laporte et al., 2003), the time consumption for both lower limbs reconstruction is an hour. For the distal femur reconstruction, the mean and 2 RMS are 1 and 2.8 mm, respectively.

2.3.2 Automatic identification of the target bone and 3D pose initialization

Automatic 3D/2D registration methods use an automatic identification and detection of the target bone structure in 2D radiographs without any user intervention to initialize and register the 3D prior model into the target bone structure in underlying 2D radiograph. In (Yu et al., 2016), an automatic machine learning-based segmentation approach using random forest regression (Chen & Zheng, 2014) is used to identify and segment the proximal femur in 2D bi-planar radiographs. A random forest regression-based approach is used to detect landmarks and segment femoral contours in 2D bi-planar radiographs. Then, to initialize the 3D pose and isotropic scale of the proximal femur, they use part of the detected ordered landmarks to find the projections of three landmarks including the center of the femoral head (FH), the greater trochanter (GT), and lesser trochanter (LT) in both frontal and lateral 2D radiographs (Figure 2.2). A triangulation-based method is used to find the 3D position of three landmarks. Given the 3D position of the corresponding three landmarks on the volumetric 3D template of the proximal femur, the similarity transformation is computed between two sets of three landmarks. Then, the 3D pose of the volumetric template is initialized in the reference space of the patient's 2D bi-planar radiographs. The mean of shape accuracy for the proximal femur is 1.3 mm which is lower than 1.5 mm reported by a relevant semi-automatic method in (Yu, Zysset, & Zheng, 2015).

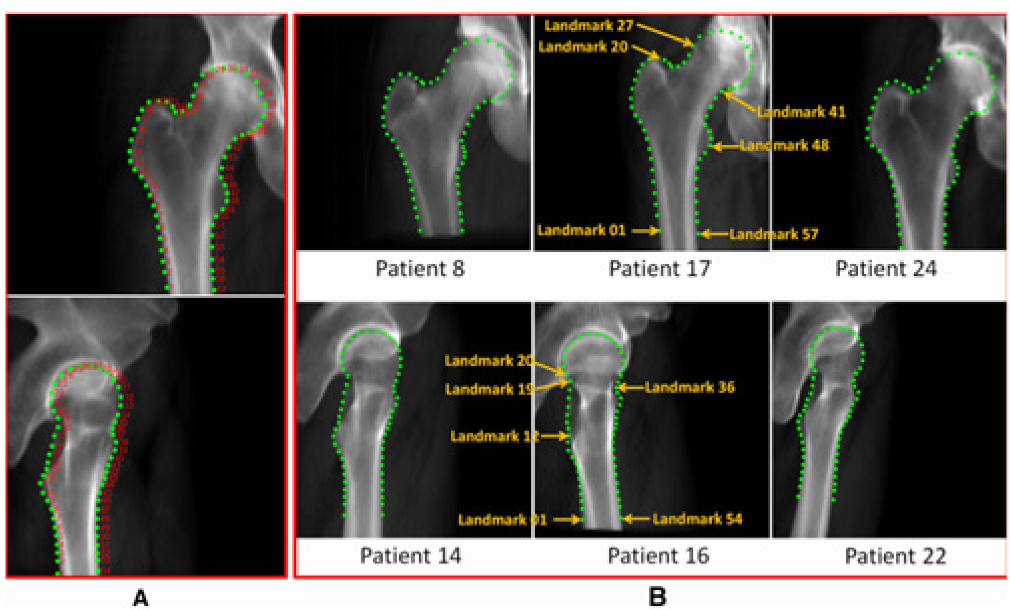


Figure 2.2 (A) Extracted ordered landmarks on frontal and lateral radiograph
(B) Top and bottom: landmarks from 27 to 41 on frontal and from 20 to 36 on lateral
Taken from Yu et al. (2016)

As another example, in (Aubert et al., 2019), for 3D spine reconstruction, the 3D pose of the center of each vertebra is automatically initialized using CNN-based landmark detection (Figure 2.3). The 3D spine reconstruction time is 34 seconds, which shows improvement in comparison to the relevant previous semi-automatic spine reconstruction methods. Regarding improvement on reproducibility, 89% of clinical measurements for spine are inside the expert's confidence intervals.

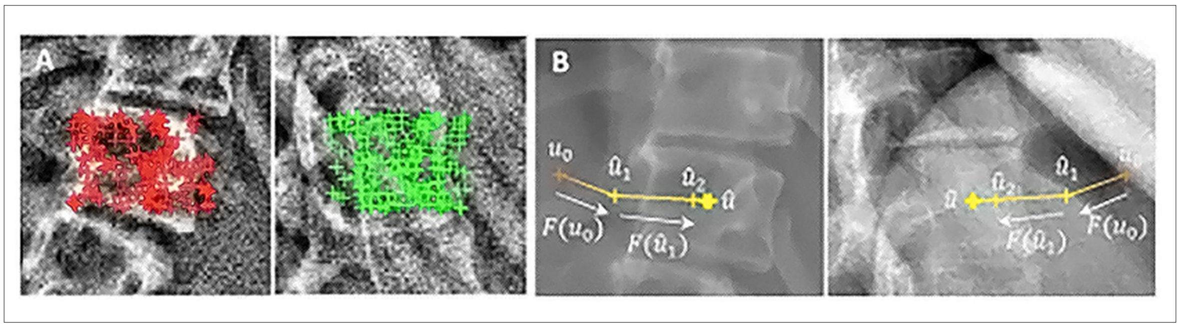


Figure 2.3 (A) Patch center sampling on frontal and lateral patches. (B) Optimizing patch center position by estimating 2D displacements of stereo corresponding points Taken from Aubert et al. (2019)

2.3.3 Conclusions on identification of target bone and 3D pose initialization

As a summary, in evolution of 3D bone model reconstruction from manual toward fully automatic methods, 3D pose initialization of the 3D prior model, as first step in 3D bone reconstruction, plays an important role in the process automation and the 3D bone reconstruction time optimization (Aubert et al., 2019 ; Yu et al., 2016). Semi-automatic 3D pose estimation methods require the manual intervention of an operator for pose initialization of the 3D model to 2D radiographs images (Baka et al., 2011, 2012 ; Laporte et al., 2003 ; Mitton et al., 2000 ; Youn et al., 2017) . These manual interventions impact the accuracy, time efficiency and reproducibility of the different methods (Chaibi et al., 2012 ; Laporte et al., 2003). Automatic 3D pose estimation methods make efforts to eliminate operator interventions in order to removing the effects of the user skills on the accuracy of the initialized 3D pose (Yu et al., 2016), decreasing the inter-reproducibility errors (Aubert et al., 2019), and 3D bone reconstruction time (Aubert et al., 2019).

2.3.4 Geometric feature-based 3D/2D rigid or similarity registration

Geometric or feature-based 3D/2D rigid or similarity registration methods define the matching term based on the geometric information in the moving(M) and the reference image(F), (Hosseinian & Arefi, 2015 ; Reyneke et al., 2019). Feature-based methods make efforts to find correspondences between identified and detected 2D geometrical features, i.e. points, contours, edges, or curves, via the above mentioned identification methods in 2D image as reference and 3D silhouette of the template model or landmark points, as moving or source model (Hosseinian & Arefi, 2015 ; Reyneke et al., 2019). Geometric feature-based methods find optimal transformation parameters (translations, rotations, and isotropic scaling parameters) by minimizing the distance between two correspondences in the source and reference (Baka et al., 2011, 2012 ; Laporte et al., 2003 ; Youn et al., 2017). Table 2.1 summarizes the reported quantitative results of geometric feature-based 3D/2D rigid or similarity registration by above-mentioned state-of-the-art methods. In (P. Gamage, Xie Q., Delmas, Xu, & Mukherjee, 2008 ; Laporte et al., 2003), an operator manually initialized the 3D generic model in 2D reference space close to optimal position, so the accuracy of the 3D pose estimation depends to the operator skills and these methods suffer from the lack of reproducibility. In (Baka et al., 2011, 2012, 2014), an operator manually initialized the mean model close to the optimal position. Hence, these methods suffer from operator interventions and lack of the reproducibility. In Table 2.1, the semi-automatic methods that use NSCC (Laporte et al., 2003) and 2D contours (P. Gamage et al., 2008) as geometric features for registration have small capture ranges for translation and rotation. (Baka et al., 2011, 2012) improved the accuracy and capture range of 3D pose estimation in comparison to (P. Gamage et al., 2008). However, (Baka et al., 2011, 2012), also, suffer from manual initialization of the SSM close to the target femur.

Table 2.1 Geometric feature-based 3D/2D rigid or similarity registration

Author & Year	Geometric Feature	Initialization	Capture Range	Precision
Laporte et al., 2003	NSCC	Manual Initialization by operator	In diaphysal axis: Translation=$[-1, +1]$ mm Rotation=$[-5, +5]°$	3D pose tracking performance = Not mentioned
Gamage et al., 2008	2D contours	Manual Initialization by operator	Starting position = 5 mm Starting orientation = 10 degrees	Maximum misalignment Error: t_z=0.96 mm r_α=0.81°
Baka et al., 2012	Optimizing a matching term based on the 2D Euclidean distance and (L2-norm) angle distance between 2D projected bone contours and X-ray edges	Manual Initialization by operator at the middle of FOV, sufficiently close to target region	$(t_x, t_y, t_z) = [-10, +10]$ mm $(r_\theta, r_\alpha, r_\beta) = [-10, +10]°$	(Mean $\pm$ STD) t_x= 0.18$\pm$1.53 mm t_y= -0.02 $\pm$0.93 mm t_z= -0.41 $\pm$1.17 mm r_θ= 0.04 $\pm$1.46 ° r_α= 0.04 $\pm$1.18 ° r_β= -1.03 $\pm$2.51 ° (r_β suffer from low accuracy)

Author & Year	Geometric Feature	Initialization	Capture Range	Precision
Baka et al., 2014	The Euclidean distance between two Gaussian Mixture Models (GMM)	Manual initialization by operator	$(t_x, t_y, t_z) = [-9, +9]$ mm $(r_\theta) = [-9, +9]°$	Success rate = 81% (2mm) Median L2 distance between ground truth and estimated landmark = 1.06 mm 3D pose tracking performance = Not mentioned

2.3.5 Iconic-based 3D/2D rigid or similarity registration

The second group of the 3D/2D rigid or similarity registration methods, iconic-based methods, exploits intensity information and defines an intensity-based matching term as a matching criterion to register source image to reference image (Zollei, Grimson, Norbash, & Wells, 2001). In the iconic-based registration, the source image is mostly digital reconstructed radiographs (DRRs) of the 3D prior model (Markelj, Tomaževič, Pernuš, & Likar, 2012) , (Figure 2.4). Iconic methods evaluate the alignment between images based on an intensity-based measure and integrate intensity information of source (M) and reference image (F) as similarity criterion in the data (dis)similarity term of the objective function of image registration, $D(U)$. In the case of assuming similar intensity values corresponding to the same body structures, different matching criteria might be used to define data (dis)similarity term, $D(U)$, such as the sum of squared difference (SSD), the sum of absolute difference (SAD), mutual information (MI), normalized mutual information (NMI), cross correlation (CCor), and correlation coefficient (CCoef) (Hatt, Speidel, & Raval, 2015 ;

Knaan & Joskowicz, 2003 ; Yao & Taylor, 2003). The iconic-based 3D/2D rigid or similarity registration methods do not need manual annotations and finding correspondences between the source and reference image (Hatt et al., 2015 ; Knaan & Joskowicz, 2003 ; Yao & Taylor, 2003). Therefore, in comparison to the geometric feature-based methods, the iconic-based methods reduce the registration process time and steps (Miao, Wang, & Liao, 2016).

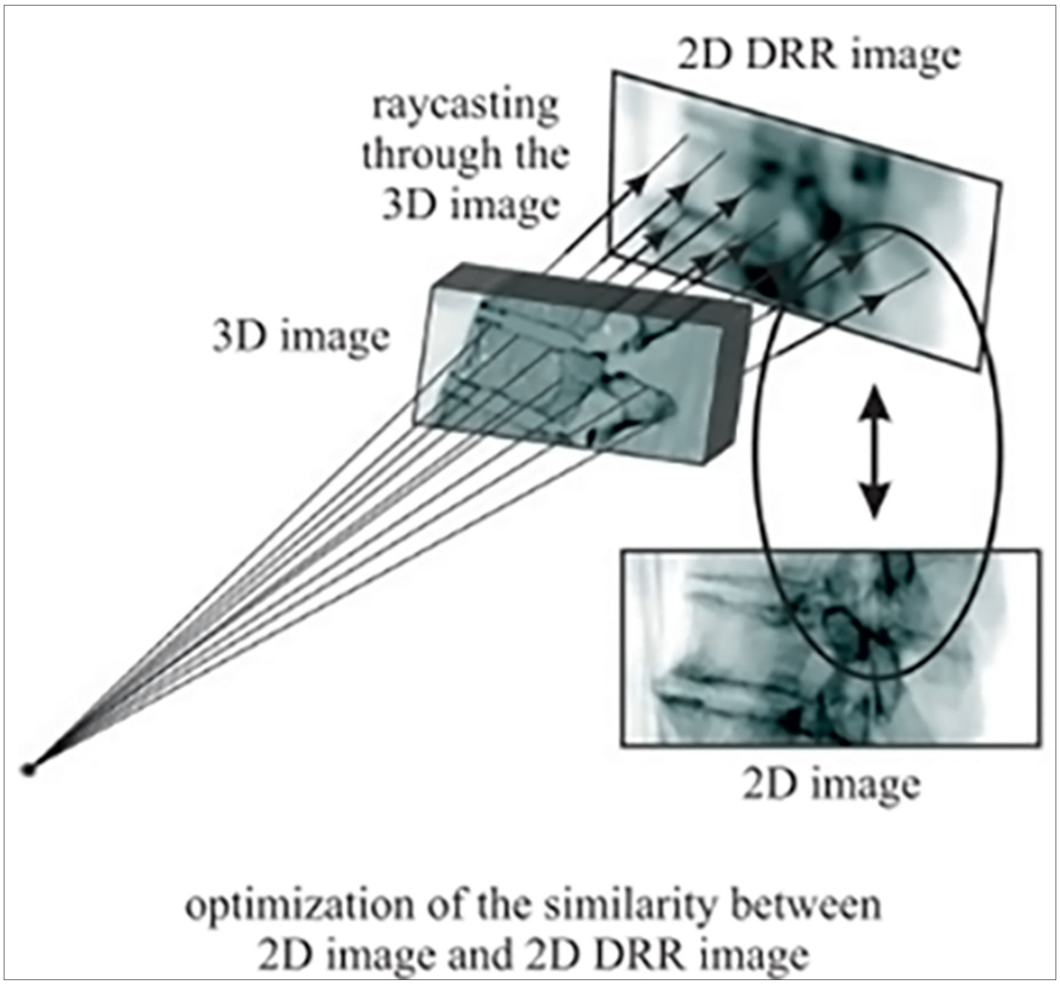


Figure 2.4 Illustration of iconic-based registration between DRR and image
Taken from Likar, Markelj, & Tomaz (2012)

Table 2.2 summarizes the reported quantitative results of iconic-based 3D/2D registration by above-mentioned state-of-the-art methods. In comparison to geometric feature-based registration methods, the proposed intensity-based optimization methods improved the capture range in both translation and rotation. However, the MI, NCC, and CC-based optimization methods are non-convex and suffer from getting trap in local optima. In Table 2.1, the proposed methods avoid local optimal by limiting the capture and manual 3D pose initialization close to optimal position.

Table 2.2 Iconic-based 3D/2D rigid or similarity registration

Author & Year	**Iconic-based Registration**	**Initialization**	**Capture Range**	**Precision**
Knaan et al., 2003	Normalized Cross Correlation (NCC) DRR-based registration	Manual initialization	$t_x = 40$ mm	Success rate ≅ 95(%) for (<2mm) mTRE = 1.1 mm
Yao & Taylor, 2003	Mutual Information (MI) DRR-based registration	Manual initialization	Not mentioned	Volume overlap = 90.1 (%)
Mahfouz et al., 2003	1-Cross Correlation Between pixel values 2-Cross Correlation between two edges (Overlapping of edges)	Manual initialization	$(t_x, t_y, t_z) =$ 16 mm $(r_\theta) = 16°$	Success rate ≅50(%) for (<5mm) (t_x, t_y) = 0.075 mm (t_z) = 0.936 mm (r_θ) = 0.347°

Author & Year	Iconic-based Registration	Initialization	Capture Range	Precision
Hatt et al., 2015	Cross Correlation (CC) DRR-based registration	Manual initialization	$(t_x, t_y) =$ $[-1.5, +1.5]$ mm $t_z =$ $[-2.5, +2.5]$ mm $(r_\theta) = [-3, +3]°$ (r_α, r_β) $= [-15, +15]°$	Success rate ≅ 97(%) for (<5mm) mTRE ≅ 1.2 mm

2.3.6 Hybrid-based 3D/2D rigid or similarity registration

Hybrid-based registration methods make efforts to benefit from the advantage points of two above mentioned categories, geometric and iconic registration methods. In this kind of registration, rigid registration is decomposed into two steps, coarse and fine step that works independently. For instance, in (Yu et al., 2016), the coarse stage is used for 3D volumetric template initialization. In the fine step, intensity-based scaled-rigid 2D-3D registration, U_{Pose}, finely aligns 2D DRRs of 3D floating volumetric template (*M*) with corresponding 2D bi-planar Radiographs, (*F*), (Yu et al., 2016). Table 2.3 summarizes the reported quantitative results of hybrid-based 3D/2D registration by state-of-the-art methods. Thanks to the hybrid-based and the coarse-to-fine registration strategy, where the intensity and geometric features in image are used, we can improve the 3D/2D rigid or similarity registration accuracy in two steps, as (Yu et al., 2016 ; Zheng, Miao, Jane Wang, & Liao, 2018). The geometric feature-based registration errors of the coarse step are enhanced using the intensity information in the fine step of the registration. The geometric features are used for the target bone identification and the 3D pose initialization and the intensity features in the image are used to improve the initialized 3D pose (Yu et al., 2016 ; Zheng et al., 2018). In 3D model-based orthopedic applications, where the prior 3D model is far from the target position, using hybrid feature-based registration and coarse-to-fine strategy helps to overcome some limitations such as: manual initialization (Yu et al., 2016), increasing capture range, and improving the registration accuracy (Zheng et al., 2018).

Table 2.3 Hybrid-based 3D/2D rigid or similarity registration

Author & Year	**Hybrid-based Registration**	**Initialization**	**Capture Range**	**Precision**
Yu et al., 2016	Control point and DRR-based registration using Mutual Information (MI)	Automatic landmark position detection	Not mentioned	3D pose Tracking performance = not mentioned
Zheng et al., 2018	Marginal space detection and DRR-based registration using Gradient Correlation (GC)	Automatic feature detection	Initial TRE = 60 mm in depth	Success rate = 93.4% (<2.5 mm) mTRE = 0.86 mm Median TRE = 0.75 mm

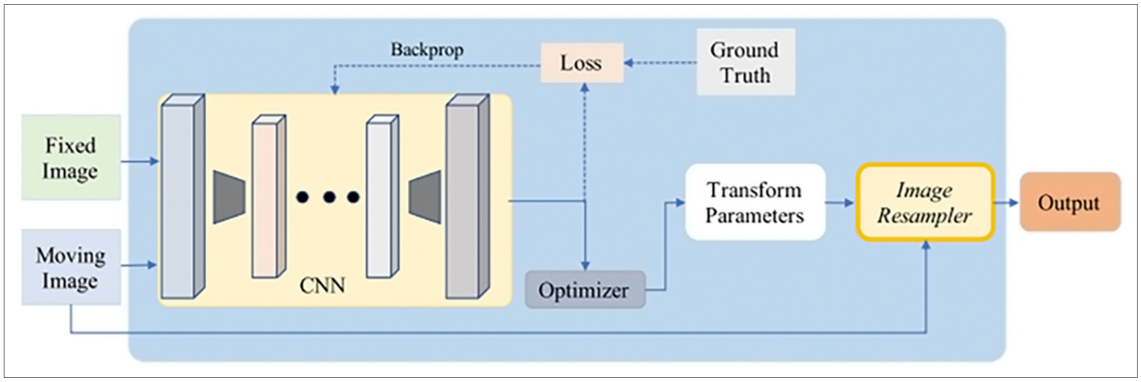


Figure 2.5 Deep leaning in an iconic-based registration framework
Taken from Haskins, Kruger, & Yan (2020)

2.3.7 Deep learning-based 3D/2D rigid or similarity registration

Recently, emerging machine learning (ML) and deep learning (DL)-based approaches (Figure 2.5) have significantly contributed to reducing operator interventions and manual initialization to facilitate the automation of 3D/2D registration (Haskins et al., 2020). The work in (Yu et al., 2016) proposes an automatic 3D/2D similarity registration of the proximal femur by using an automatic machine leaning-based segmentation, via random forest regression, of the target bone structure. As an advantage, they initialize a template 3D model without user interactions and remove manual initialization. In addition, using automatic initialization increases reproducibility of the work, since the registration results do not depend on operator skills. However, they need to collect a training data set to train the ML-based model. In an automatic 3D/2D rigid registration of a transesophageal echocardiography (TEE) probe, (Zheng et al., 2018) propose an automatic 3D pose initialization via the marginal space learning method. The deep convolutional neural network (CNN) performs very highly in learning the non-linearity of the mapping function between input image features and transformation parameters (Haskins et al., 2020 ; Liao et al., 2016 ; Miao et al., 2016 ; Zheng et al., 2018).

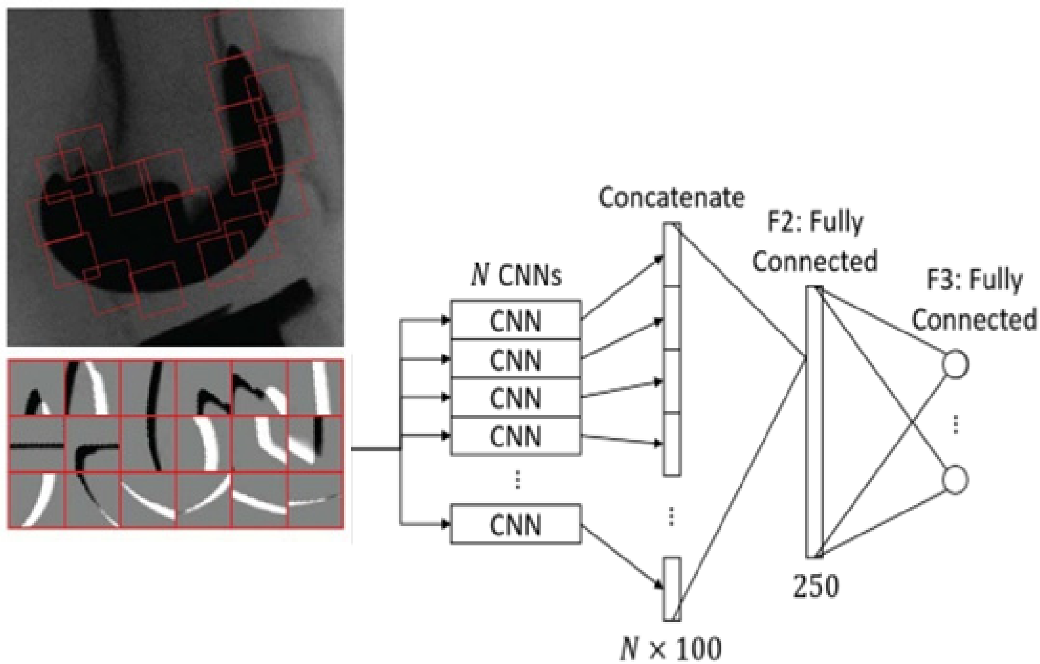


Figure 2.6 (Left) Input local intensity residuals (LIRs). (Right) CNNs-based regression model to estimate transformation parameters residuals
Reproduced and adapted from Miao et al. (2016)

In contrast to (Yao & Taylor, 2003 ; Yu et al., 2016), a deep learning-based optimization method (Liao et al., 2016 ; Miao et al., 2016 ; Zheng et al., 2018) overcomes limitations of non- convex intensity-based (dis)similarity term optimization by using a better prior, i.e., a large regularized model in a high dimensional space, to guide the optimization process and get rid of local optimums. In deep learning, by increasing the parameters of the model, an escape route might appears in those extra dimensions, which makes it possible for the model to move out of the local optimums. In (Miao et al., 2016), known rigid 3D objects such as the knee prosthesis 3D model, the trans-esophageal echocardiography probe (TEE), and a tooth implant are registered into the patient's 2D frontal radiograph. Hierarchical regression models (Figure 2.6) based on a deep convolutional neural network (CNN) are trained to regress six rigid transformation parameters (6DOF) from input images. The input images are local intensity residuals (LIRs), which are intensity differences between Digitally Reconstructed Radiograph (DRR) of a known rigid 3D object, as source, and patient's radiograph, as reference image. LIRs are computed in extracted local patches centered at 2D projection of 3D pose index points defined on the 3D object, (Figure 2.6 (Left)), (Miao et al., 2016). The proposed deep learning-based registration method shows enough robustness to reproduce the same registration results from different starting positions by the root mean squared distance in the projection direction (RMSDproj) equal to 0.005 mm (Miao et al., 2016). Instead of iteratively optimizing 3D rigid transformation parameters, $U_{Pose} = G_{3D} = \{t_x, t_y, t_z, r_\theta, r_\alpha, r_\beta\}$, over a non-convex data similarity measurement term, $D(U_{Pose}) = D(G_{3D}) = D(F, M, G_{3D})$, as:

$$\frac{\partial(E(U))}{\partial(G_{3D})} = \frac{\partial(E(U))}{\partial(D(F, M, G_{3D}))} \cdot \frac{\partial(D(F, M, G_{3D}))}{\partial(G_{3D})} \tag{2.5}$$

They present convolutional neural networks CNNs-based regressor models to directly estimate 3D rigid transformation parameters from image intensity residuals (Miao et al., 2016). The following objective function which is a loss function, Eq. (2.6), is minimized during the training the CNNs. During the training, W which is the vector of weights is learned to find optimal target 3D rigid transformation parameters, G_{3D}, and is updated, i.e., $\frac{\partial(D(U_{Pose}))}{\partial(W)}$, via back propagation training method in conjunction with stochastic gradient descent (SGD) optimizer. The objective function to minimize (dis)similarity is defined as a Euclidean loss function (Miao et al., 2016):

$$D(F, M, G_{3D}) = \frac{1}{K} \sum_{i=1}^{K} \|F_i - U_{Pose}(M_i; W)\|_2^2 \tag{2.6}$$

Here, K is the number of training samples. The last trained value of vector W, learned weigths, is used to estimate transformation parameters for unseen pairs of input images (Miao et al., 2016). In the most relevant application, the total knee arthroplasty (TKA) 3D kinematics, the Mean Target Registration Error (mTRE) is 6.73 mm in 3D, (Miao et al., 2016). In another study, (Zheng et al., 2018) propose a coarse-to-fine strategy for 3D/2D registration. Rather than manual initialization of the known prior 3D model and then CNN-based registration in (Miao et al., 2016), they use marginal space learning method for automatic initialization of the know prior 3D model. Then, to increase the capture range, they use CNN-based coarse and then CNN-based fine registration, using intensity features (Zheng et al., 2018). They estimate transformation parameters using CNNs-based regressors from input intensity residuals between DRRs and X-ray image. The Mean Target Registration Error (mTRE) is 1.18 mm in 3D. The Root Mean Square Error (RMSE) of two out-plane rotations is equal to 4.59° (Zheng et al., 2018). In the same application, transesophageal echocardiography (TEE) probe tracking, Table 2.4 shows the positive impacts of using coarse-to-fine strategy of 3D/2D registration of in comparison to (Miao et al., 2016). They increased the capture range of the rigid 3D/2D registration (Table 2.4) using the coarse-to-fine technique in comparison with (Miao et al., 2016 ; Zheng et al., 2018).

Table 2.4 Comparison of translation and two rotations' capture ranges

	Translation in depth (mm)	Out-plane rotation around X-axis (°)	Out-plane rotation around Y-axis (°)
(Miao et al., 2016)	$[0, +7.23]$	$[-15, +15]$	$[-15, +15]$
(Zheng et al., 2018)	$[0, +60]$	$[-45, +45]$	$[-90, +90]$

2.3.8 Conclusions on 3D pose estimation

The first step, in which an anatomical template model is positioned in the radiograph reference space, can be performed via a manual (Baka et al., 2011 ; Youn et al., 2017), semi-automatic (Chaibi et al., 2012), or automatic (Yu et al., 2016) identification and detection of the target bone structures in 2D radiographs, and then by the 3D/2D rigid or similarity registration. In this first step, accurate 3D pose and isotropic scaling parameters along with a large capture range for out-of-plane rotations could significantly contribute to achieve a good fitting in 3D bone reconstruction (Baka et al., 2011 ; Chaibi et al., 2012). Most studies, however, use a manual initialization of the object pose close to the optimal position (Baka et al., 2011 ; Chaibi et al., 2012 ; Youn et al., 2017). An automatic estimation of accurate 3D pose parameters and isotropic scaling of the bone structure remains a challenging task in 3D/2D similarity registration (Youn et al., 2017 ; Yu et al., 2016). Among the reported accuracies of 3D/2D registrations in relevant 3D model-based clinical applications, (Baka et al., 2011) obtained one of the best results, which proposed a semi-automatic 3D/2D geometrical feature-based similarity registration of the distal femur into bi-planar stereo X-rays. The Mean Absolute Errors (MAE) of the three translations, the three rotations, and the isotropic scaling in 3D are equal to 0.86 mm, 1.30°, and 1.55(%), respectively (Baka et al., 2011). (Hatt et al., 2015) proposed a semi-automatic 3D/2D rigid registration method based on volume gradients using a computed image similarity metric for transesophageal echocardiography (TEE) probe tracking with five degrees of freedom (5DOF). The mean target registration error (mTRE) is 1.2 mm in 3D (Hatt et al., 2015). Recent developments in Convolutional Neural Networks (CNN) for 3D/2D registration addressed major limitations like small capture ranges (Table 2.4), and significantly improved robustness of the 3D pose estimation (Miao et al., 2016 ; Zheng et al., 2018). (Miao et al., 2016), after manual 3D pose initialization, employed a CNN-based regressor to directly estimate six rigid transformation parameters (6DOF) to fit rigid objects such as a 3D model of the knee prosthesis, a transesophageal echocardiography (TEE) probe, or an implant into one X-ray image. To remove the need of manual pose initialization and significantly reduce the limitation of the small capture ranges (Table 2.4), (Zheng et al., 2018) propose a fully automatic monoplane a coarse-to-fine CNN-based TEE probe tracking using an automatic initialization by the marginal space learning method. Some drawbacks limit the above-mentioned applications for personalized 3D bone reconstruction such as:

1) Previous methods estimate six rigid transformation parameters (6DOF) because the object is known;
2) They suffer from a lack of sub-millimeter, sub-degree, and sub-percentage precision (Baka et al., 2011 ; Hatt et al., 2015 ; Miao et al., 2016 ; Zheng et al., 2018), particularly in the proximal distal rotation axis (Baka et al., 2011 ; Zheng et al., 2018);

3) Most of them need a manual initialization.

However, in 3D model-based orthopedic applications, using automatic initialization (Yu et al., 2016 ; Zheng et al., 2018) and a coarse-to-fine CNN-based strategy (Zheng et al., 2018) helps to remove operator intervention, operator skill dependency, manual initialization close to optimal position, and increase capture range of rigid or similarity registration .

2.4 3D/2D non-rigid registration

The 3D model reconstruction from 2D bi-planar radiographs consists of establishing a geometric relationship between a prior 3D model, as a source model, and 2D radiographs, as reference images, using firstly 3D/2D rigid or similarity registration, and secondly, 3D/2D non-rigid registration techniques (Baka et al., 2011 ; Youn et al., 2017 ; Yu et al., 2016). In 3D/2D non-rigid registration, deformable prior 3D model selection and regularization term, based on the expected solution for 3D shape deformation algorithm and the type of anatomical organs, drive 3D shape deformation of moving prior 3D model (M) to 2D target images (F) (Lisa Y. W. Tang and Ghassan Hamarneh, 2013). In the objective function of 3D/2D registration, Eq. (2.1), the regularization term, $R_i(U)$, enforces specific constrains based on the type of deformation model and expected solution to restrict the type of U in order to obtain an optimal transformation. The definition of regularization term depends on the kind of 3D shape deformation methods such as interpolation-based and knowledge-based deformation methods (Lisa Y. W. Tang and Ghassan Hamarneh, 2013 ; Sotiras, Davatzikos, & Paragios, 2013). The regularization term by constraining the 3D shape deformation provides inverse consistency, symmetry, topology preserving, and diffeomorphism. Inverse consistency property constrains forward and backward transformations in order to be the inverse mapping of each other. The symmetry characteristic penalizes asymmetric registration to estimate the inverse transformation. Topology preservation criteria play a key role in restricting the transformation algorithm to be one-to-one instead of crossing in order to obtain inevitability for topology preserving mapping (Sotiras et al., 2013). Diffeomorphism criteria, to preserve the topology of the 3D deformable model, require both smoothness and inevitability, being bijective, continuous, and its inverse is bijective (Lisa Y. W. Tang and Ghassan Hamarneh, 2013), so transformation function and its inverse are differentiable (Sotiras et al., 2013). As a reminder, 3D/2D registration combines both global transformation (3D pose estimation) and local 3D shape and 3D scale deformation (3D/2D non-rigid registration) in such a way that global transformation is followed by local shape deformation, data (dis)similarity term (matching term) is defined as:

$$U_{Shape}(U_{Pose}(M)) = D\,(F, M, LoG_{3D}) \tag{2.7}$$

Where, U_{Shape} is local 3D shape and 3D scale deformation (3D/2D non-rigid registration), U_{Pose} is 3D pose estimation, and $D\,(F, M, LoG_{3D})$ is data (dis)similarity term between M and F after 3D/2D registration. Based on the reviews (Goswami & Kr., 2015 ; Hosseinian & Arefi, 2015 ; Reyneke et al., 2019), in the lower limb 3D bone reconstruction, state-of-the-arts propose various methods for 3D/2D non-rigid registration.

2.4.1 Prior 3D model presentation

In 3D model reconstruction, a known prior 3D model such as a CT-scan-based 3D model (Fang et al., 2020), a statistical shape model (SSM) (Baka et al., 2011, 2012 ; Youn et al., 2017 ; Yu et al., 2016 ; Yu, Tannast, & Zheng, 2017), or a generic template model (Chaibi et al., 2012 ; Khameneh, Vazquez, Cresson, Lavoie, & de Guise, 2021 ; Yu et al., 2016) could be registered into the patient's calibrated 2D bi-planar radiographs. The choice of the 3D prior model drives the 3D bone reconstruction process.

2.4.1.1 CT-scan-based 3D model

In 3D model-based orthopedic applications, 3D imaging acquisition is used to construct personalized 3D model. Figure 2.7 shows reconstructed 3D models from segmented 3D CT-scans.

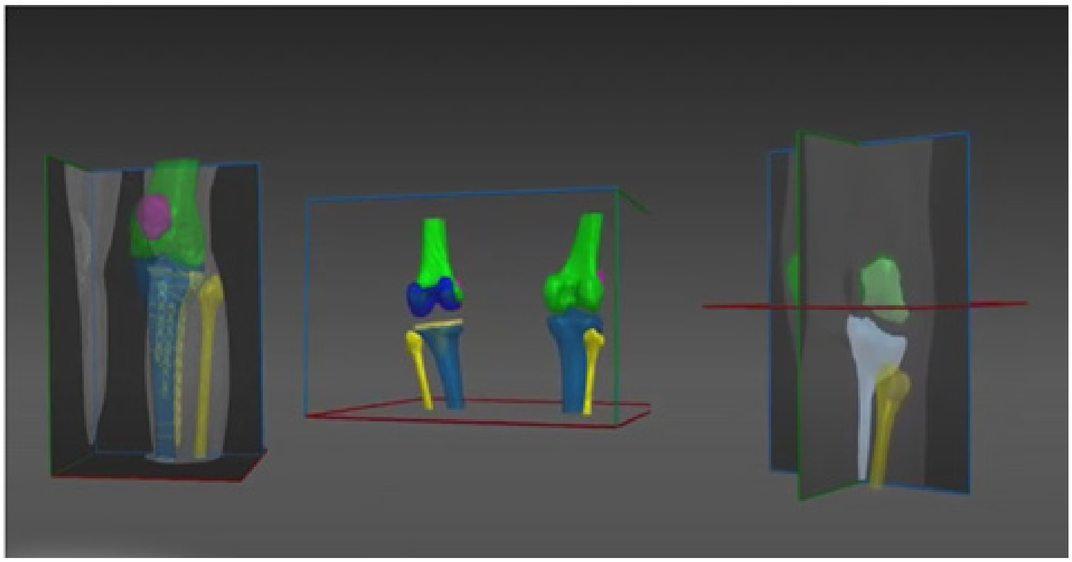


Figure 2.7 (left & right) CT-scan-based segmented distal femur and proximal tibia. (Middle) patient-specific CT-scan-based 3D model
Taken from https://www.rsipvision.com/bones-skeleton-segmentation

For instance, in the laboratory LIO, one of the most popular software to manually segment the femur in 2D slices of the 3D CT-scans is the SliceOmatic software tool (Figure 2.8). All 2D slices of the stack are manually segmented and saved as .tag files. Then, all .tag files are read to a volume. Afterwards, in Matlab software, by using isosurface(.) function, a 3D surface mesh is reconstructed based on the volume.

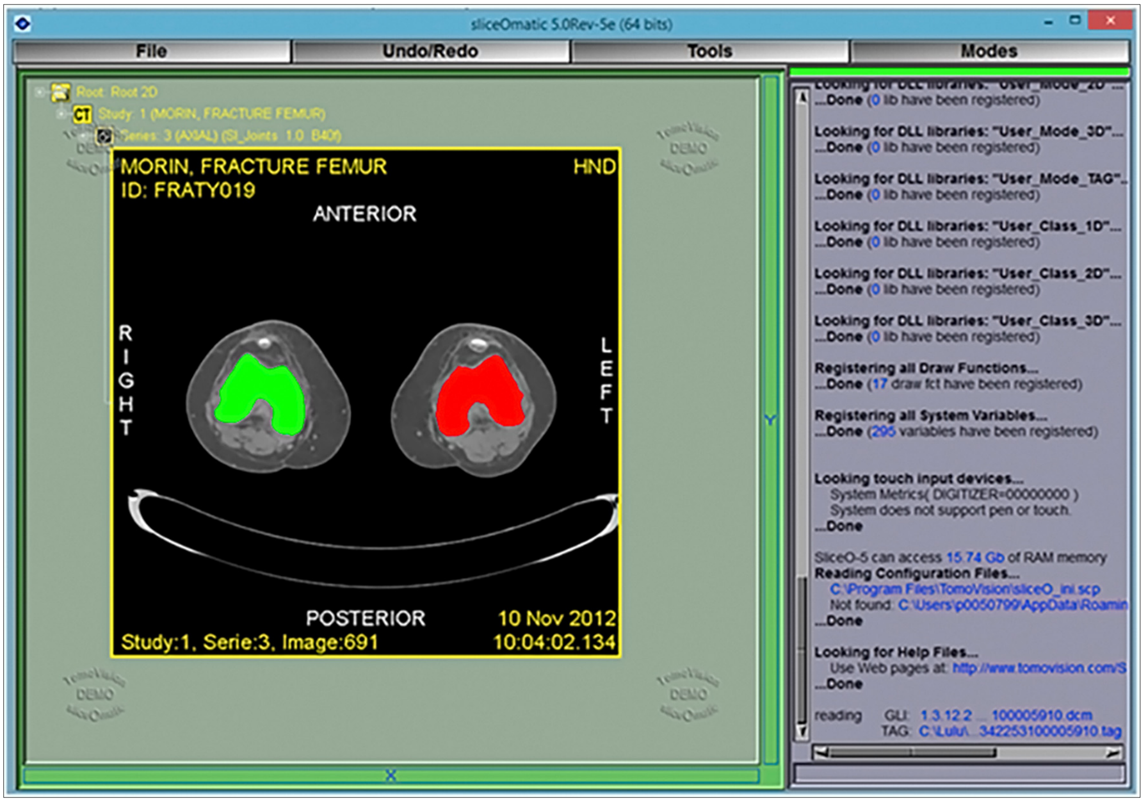


Figure 2.8 Illustration of SliceOmatic software tool to segment the femur on 2D slice of 3D CT-scans

In 3D model-based application, (Zollei et al., 2001) register prior 3D model from segmented CT-scans of the patient into 2D radiographs using DRR-based and mutual information-based registration method. As another example, in 3D shape reconstruction of the lumbar vertebra, (Fang et al., 2020) register CT-volume prior 3D model into 2D radiographs using DRR-based and mutual information-based optimization. 3D CT-scans operate with patients in reclining position. However, natural standing position of the patients is very important in 3D femur shape deformation analysis and the clinical 3D geometric parameter measurements (Chaibi et al., 2012). Moreover, 3D CT-scans provide high level of radiation dose and expensive treatments (Goswami & Kr., 2015 ; Hosseinian & Arefi, 2015 ; Reyneke et al., 2019).

2.4.1.2 Statistical shape model (SSM)

In the personalized 3D model reconstruction, to avoid 3D imaging and the high levels of radiation doses and costs related to 3D CT-scans acquisitions, the state-of-the art methods propose to use statistical shape model (SSM) of the bone structures (Figure 2.9) as source model to register into 2D radiographs (Aubert et al., 2019 ; Baka et al., 2011, 2012 ; Fleute & Lavallée, 1999 ; Yao & Taylor, 2003 ; Youn et al., 2017).

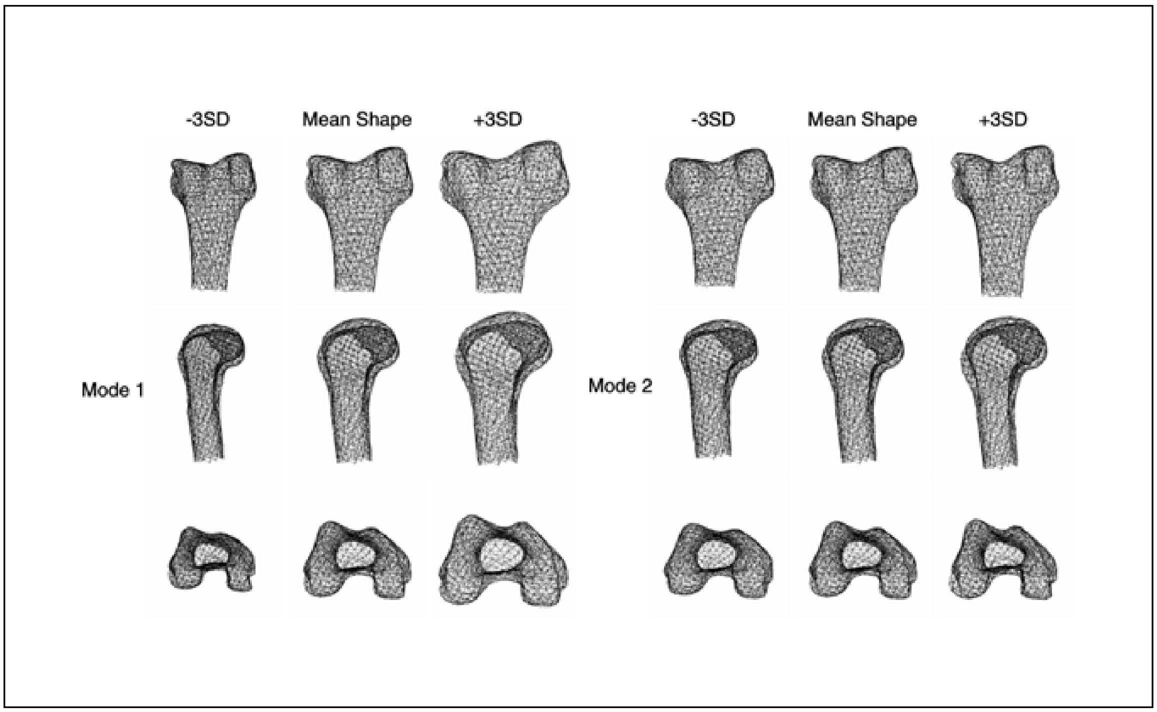


Figure 2.9 Illustration of mean shape and standard deviations of two deformation modes
Taken from Fleute & Lavallée (1999)

In 3D/2D non-rigid registration, there are several methods to create SSM as prior 3D model, which is a deformable shape template. (Cootes, Taylor, Cooper, & Graham, 1995) introduce a statistical shape model by incorporating statistical prior knowledge about anatomical variability among a patient population. In this approach, the point distribution model is built from N training shapes segmented from CT data. A correspondence algorithm is applied to represent each shape (M) with the same number n of corresponding landmark points $F_j, j =$

$1, \ldots, n$, which are interconnected to form triangulated surfaces (Dubousset et al., 2007 ; Ferrarini et al., 2007). By applying Principal component analysis (PCA) on the deviations of the mean shape, we obtain a model of the form as Eq. (2.8), (Cootes et al., 1995):

$$x \approx \bar{x} + \Phi b_x \tag{2.8}$$

Where $\bar{x}$ is the mean shape, Φ is the matrix of the main modes of variation, and b_x is the representation of the shape x in parameter space. In SSM-based 3D femur reconstruction, Principal Component Analysis (PCA) is a very useful and popular method used to handle and regularize 3D shape deformation (Baka et al., 2011, 2012 ; Fleute & Lavallée, 1999 ; Yu et al., 2016, 2015). Following (Cootes et al., 1995), in the distal femur reconstruction application, (Fleute & Lavallée, 1999) uses SSM and define modes of shape variation using PCA as Eq. (2.9) :

$$m = \overline{m} + \sum_{i=1}^{t} w_i e_i , \quad 1 \leq t \leq 3 \tag{2.9}$$

Where m, is a mode of shape variation, $\overline{m}$ is the mean shape, e_i is eigenvectors of mode variations, w_i is the weighting factor for the i^{th} variation vector. t is limited to only 3 variations to ensure only the important deformations are extracted. The comparison between Eq. (2.8) and (2.9) shows that Eq. (2.9) uses the three most important weighted eigenvectors to represent the new instance of the shape and create a new mode variation. (Fleute & Lavallée, 1999) uses 10 dry femurs to construct the mean shape. In (Baka et al., 2011), to construct a SSM model of the distal femur as a prior 3D shape, 43 CT-based reconstructed 3D models of the femur are used to prepare a training population. Then, the PCA method is used to obtain the statistical shape model. In this way, the geometrical shape information and its variations across the training shape population are used to build a prior shape template (Fleute & Lavallée, 1999). Although the SSM is used in many 3D bone reconstruction applications, but an SSM is a variation of the mean shape limited to the training shapes' variations and knowledge, which may not be complete enough to create a perfect match to the reference images (Cresson et al., 2010). However, SSM helps to avoid CT-scans acquisition to construct a personalized 3D model for each patient. Using SSM restricts the 3D reconstruction to plausible forms (Baka et al., 2012). To construct a SSM, preparing a training set of the CT-scan-based 3D models is time consuming, costly.

2.4.1.3 Generic 3D model

In contrast to SSM which uses statistical information, generic 3D model of the bone structure is a single geometrical surface deformable model acquired by (CT)-scans-based reconstruction from a single dry bone structure (Laporte et al., 2003). For instance, in (Laporte et al., 2003), a generic 3D model of the distal femur is generated using projection of a mesh, which includes 556 points and 1100 triangles, on the surface of a CT-scan-based reconstructed 3D model of a dry femur via SliceOmatic© software (Mitulescu, Semaan, De Guise, Leborgne, & Adamsbaum, 2001). The SliceOmatic software provides an automatic segmentation of CT-scan slices, and then manual correction of the segmented slices to obtain a CT-scan-based 3D model reconstruction (Mitulescu et al., 2001). Figure 2.10 shows a generic 3D model which is a CT-scan-based reconstructed 3D surface mesh of the femur. Unlike the SSM, to generate a generic 3D model of the bone structure, we do not need to prepare a training set of the CT-scan-based 3D models and a generic 3D model is a single template model constructed based on the CT-scans of a single bone. Hence, using a generic model leads to save time and cost, and avoids high levels of radiation exposures. In 3D femur reconstruction, the generic 3D model is registered into 2D bi-planar radiographs (Chaibi et al., 2012 ; Cresson et al., 2010 ; Laporte et al., 2003). In 3D/2D non-rigid registration, interpolation or approximation theory considers known displacement for a limited set of points defined on the generic 3D model (Chaibi et al., 2012) and it interpolates the displacement field for the rest of the points. The most famous methods in interpolation-based strategies are radial basis functions (RBFs) (Sotiras et al., 2013), elastic body splines (Sotiras et al., 2013), free-form deformations (Sotiras et al., 2013). The most common radial basis functions include thin-plate-spline model, which assumes infinite boundary to regularize objective function (Sotiras et al., 2013), discrete Fourier transform (DFT) (Lisa Y. W. Tang and Ghassan Hamarneh, 2013) and discrete cosine transformation (DCT) (Yu et al., 2016). In medical imaging application, thin-plate-spline model is commonly used (Lisa Y. W. Tang and Ghassan Hamarneh, 2013). Thin-plate-spline provides global effects, i.e. deformation in a region has impacts on other regions, but for non-rigid registration, we need local deformation, locally supported functions such as basis-spline function or its extended version such as free-form deformation model (Sotiras et al., 2013). Free form deformation, U_{Shape}, is commonly used in medical image registration, and it is a nonlinear transformation, that provides more localize deformation (Sotiras et al., 2013). It is a control-point based deformation, and a dense deformation might be provided by the summation of 3-D tensor product of 1-D cubic-B spline, as Eq. (2.10), (Sotiras et al., 2013):

$$U_{Shape}(X) = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{p=0}^{3} \beta_{l,n}(r)\, \beta_{m,n}(s)\, \beta_{p,n}(t) d_{i+l,j+m,k+n} \tag{2.10}$$

A generic 3D model, with a pre-defined set of handles, is a flexible model to be deformed and fitted into the target bone structure (Chaibi et al., 2012 ; Cresson et al., 2010). In (Chaibi et al., 2012 ; Cresson et al., 2010), using a generic 3D model helps to apply a fast 3D femur reconstruction method using a small set of MLS (Moving Least Square) handles defined on the generic 3D model, which are used for 3D shape deformation.

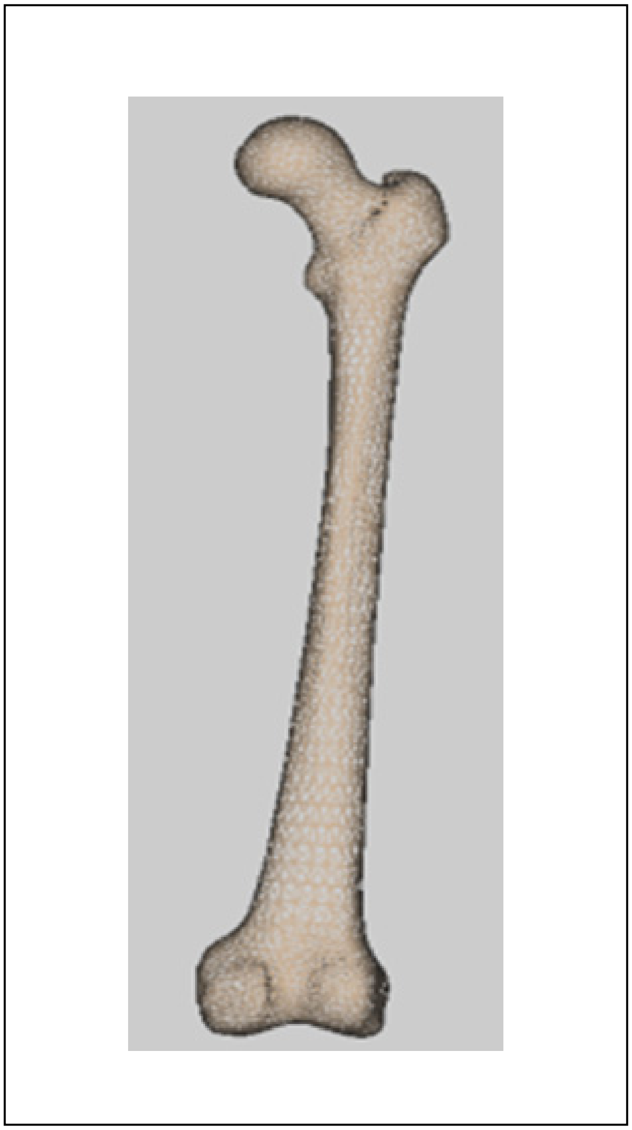


Figure 2.10 A generic 3D model of the femur

(Cresson et al., 2008; Cresson et al., 2010, Chaibi et al., 2012) use moving least-squares (MLS) deformation to interpolate 3D shape deformation. MLS deformation is based on interpolation-theory. This method interpolates new positions of the vertices on the entire model domain (*M*) based on the new positions of a set of 17 handles defined on the model domain (*M*). In this method, two sets of 3D positions of 17 handles, $P = \{p_i \in R^3, i = 1, \dots, 17\}$ and corresponding deformed positions $Q = \{q_i \in R^3, i = 1, \dots, 17\}$, are defined, both P and Q are moving handles (*M*), (Figure 2.11). Each vertex (v) on the 3D mesh of the surface (*M*) is deformed by finding the optimal transformation $\widehat{U}_{Shape}$ that minimizes:

$$\widehat{U}_{Shape} = \arg min_{U_{Shape}} \sum_i w_i D(U_{Shape}) \tag{2.11}$$

Where i is the number of the handles, and $D(U_{Shape})$ is the position error as Eq. (2.12):

$$D(U_{Shape}) = \|q_i - p_i\|^2 \tag{2.12}$$

And the weight w_i is:

$$w_i = \frac{1}{d(p_i, M)^{2\alpha}} \tag{2.13}$$

The function $d(.)$ measures the distance between each vertex (v) on the generic 3D mesh of the surface (*M*) and the handle p_i with Euclidean norm. The parameter $\alpha > 0$ controls the effect of distant handles on the deformation of the vertex (v). For each handles, the closest vertices get more weights and are affected more.

2.4.1.4 Conclusions on 3D prior model selection

In summary, for 3D/2D non-rigid registration task, state-of-the-art methods use different types of the prior 3D model. The CT-scans-based 3D model reconstruction methods suffer from the high levels of radiation dose and costs. To avoid 3D CT-scans acquisitions and high levels of radiation exposures and expensive treatments, state-of-the-art methods make efforts to use statistical shape model (SSM) (Aubert et al., 2019 ; Baka et al., 2011, 2012 ; Youn et al., 2017) or a generic 3D model (Chaibi et al., 2012 ; Cresson et al., 2010 ; Yu et al., 2016) rather than CT-scans-based 3D models. Moreover, in some cases we do not have access to acquire CT-scans of the patients. The statistical shape model constraints 3D/2D non-rigid registration (3D shape deformation) to a limited number of previous observed data in learned population, hence it needs a big database to include and learn more 3D shape variations. Moreover, both normal and pathologic samples are required in training to provide more accurate results for shape deformation (Cresson et al., 2010 ; Yu et al., 2016). Furthermore, the statistical shape model is built based on the population samples and then is registered to the patient image, so the SSM preparation is time consuming and brings about computational time for SSM-based 3D model reconstruction methods (Baka et al., 2011, 2012 ; Yu et al., 2016). Furthermore, statistical shape models are constructed based on segmented training shapes, so a manual or an automatic segmentation has its own drawbacks such as requirement of skill expert and inaccurate segmentation, respectively (Yu et al., 2016). Instead of using SSM, by using the generic 3D model and fitting it to the input images of the patient, we get rid of the training database preparation. So we can save the time for the 3D prior model generation and get less computational time (Yu et al., 2016). Unlike the SSM which is limited to the training samples, in 3D model reconstruction, using a generic 3D model with a set of pre-defined handles, which is a single template, is not restricted to the population. Using a generic 3D model with a set of handles allows to apply fast 3D shape deformation methods (Chaibi et al., 2012 ; Cresson et al., 2010).

2.4.2 SSM-based 3D/2D non-rigid registration

An statistical shape model (SSM) requires data samples to train the statistical model with high dimensionality, and the registration result is limited to the already seen deformations (Reyneke et al., 2019). To solve the problem of 3D/2D non-rigid registration, (Fleute & Lavallée, 1999) employ a statistical shape model of distal femurs for non-rigid 3D/2D registration into a few 2D X-ray views. The 3D statistical prior model (*M*) is deformed non-rigidly to the contours segmented on the 2D X-ray images (*F*). The model position and shape-parameters are iteratively updated to obtain an optimal fit $U = U_{Shape}(U_{Pose})$ by minimizing the distance, $D\left(U_{Shape}(U_{Pose})\right)$, between the contours of the model surface (*M*) and the contours formed by a discrete number of projections rays within the x-ray acquisition setting (*F*) using the ICP algorithm (Iterative closest point), (Besl, P. and McKay, 1992).

They define objective function of 3D/2D non-rigid registration based on the generalized ICP algorithm as Eq. (2.14), (Fleute & Lavallée, 1999):

$$E(U) = E(R, T, w_1, \dots, w_t) = \sum_{j=1}^{P} \min_{1 \le k \le G} \left\| F_j - (RM_k(w_1 \dots w_t) + T) \right\|^2 \quad (2.14)$$

where R (rotation) and T (translation) are rigid registration parameters, pose parameters U_{Pose}, and $(w_1 \dots w_t)$ are deformation parameters in PCA formula, Eq. (2.9). Given pose parameters, deformation parameters are optimized using the generalized ICP method. For regularization of shape deformation, they force model deformation parameters to specific bounds in order to constrain the recovered model in an anatomical reasonable range (Fleute & Lavallée, 1999). As another example, (Baka et al., 2011) proposes a method for 3D/2D shape reconstruction of the 3D distal femur surface from stereo (two or more) x-ray imaging (*F*) using statistical shape models. A statistical shape model (*M*) as a shape prior is built from N training segmented shapes derived from CT-scan data. The methodology is based on the optimization of an objective function, as Eq. (2.15), that contains both a data fitting term E_{fit} , which is a data (dis)similarity term and a shape prior term E_{prior}, as a regularization term:

$$E(U) = E_{fit} + \lambda E_{prior} \quad (2.15)$$

where λ is a weighting factor, and i=1 as they have one regularization term. λE_{prior}, as a regularization term, represents the prior knowledge of plausible shapes and regularizes shape deformation in order to keep the deformed shape close to the mean shape. To penalize inappropriate reconstructed shapes during optimization process, the Mahalanobis distance, as regularization term, is computed between the current 3D shape and the model mean, the shape prior, as Eq. (2.16):

$$E_{prior} = b_x^T \Sigma^{-1} b_x \quad (2.16)$$

where ∑ is the covariance matrix of the aligned shapes in the parameter space (Baka et al., 2011). As another example, (Yao & Taylor, 2003) studies deformable 3D/2D registration based on a statistical pelvis model and optimize an intensity-based similarity measurement. SSM-based non-rigid registration methods can use both the geometric or intensity features to fit the SSM into the 2D radiographs. As another advantage, SSM-based non-rigid registration restricts 3D shape deformation to plausible shapes and avoids undesirable distortion, so it does not need strong regularization. As limitation, SSM-based non-rigid registration methods constrain the 3D shape deformation to the training population. Hence, SSM-based methods need to prepare a big dataset to include as many as possible 3D shape variations. Preparation and training a dataset is time consuming and costly. The reported 3D reconstruction accuracy with manual initialization in (Fleute & Lavallée, 1999), (Baka et al., 2011), and (Yu et al., 2015), in comparison to ground truth CT-scans-based models, are RMS = 0.99 mm, RMS = 1.68 mm, and mean = 1.5 mm, respectively.

2.4.3 Generic 3D model-based 3D/2D non-rigid registration

In 2D radiographs-based 3D reconstruction of the femur, (Pavan Gamage et al., 2010), firstly, apply 2D non-rigid registration of edge contour data points by minimizing dis-similarity measurements between the projected contour of a generic 3D model and 2D extracted edges of the femur in radiograph. Then, a control-point-based free form deformation, Eq. (2.10) as described in section 2.4.1.3, is used to interpolate 3D shape deformation entire the 3D mesh of control points. In another survey, (Yu et al., 2016) proposes regularized deformable B-spline 2D-3D non-rigid registration framework to align 3D volumetric template, i.e., $\{I(X_M)\}$ where X_M is a point in the template volume (*M*), to 2D bi-planar X-ray images *(F)* in order to construct 3D patient-specific model. A set of 3D control points are defined on the domain of the volumetric template model (Figure 2.11). They compute control-point-based free-form deformation Eq. (2.10) as 3-D tensor product of the 1-D cubic B-spline (Yu et al., 2016). To regularize shape deformation, they apply an adaptive regularization approach (Yu et al., 2016).

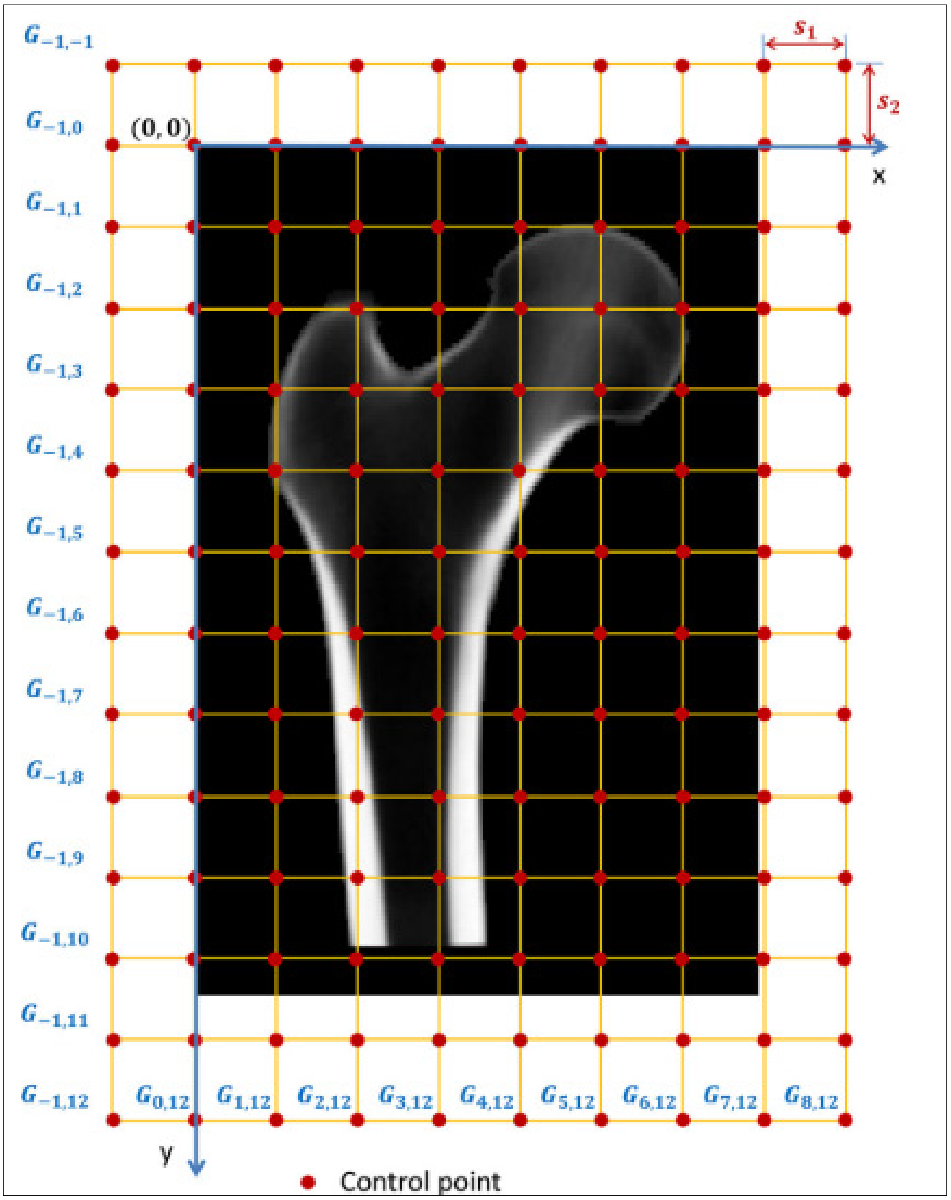


Figure 2.11 Illustration of uniform distribution of control points in FFD
Taken from Yu et al. (2016)

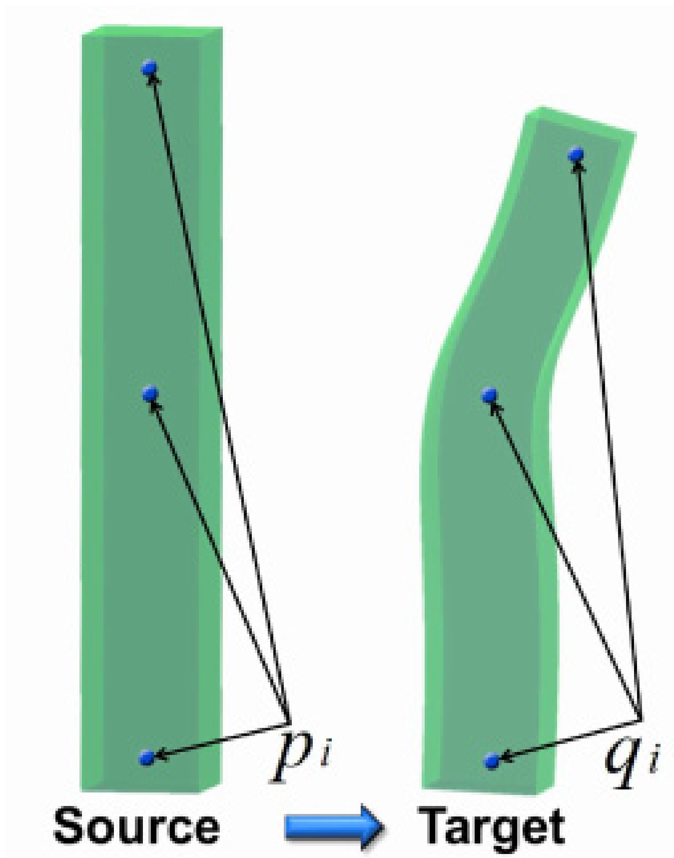


Figure 2.12 Illustration of MLS deformations between source and target model
Taken from Cresson et al. (2010)

For 3D/2D non-rigid registration, in the 3D femur reconstruction application, (Chaibi et al., 2012 ; Cresson et al., 2008, 2010) use an MLS deformation method (Figure 2.12) to register generic 3D model of the femur into the patient's EOS®2D bi-planar radiographs. In (Chaibi et al., 2012), the MLS handles are manually adjusted by a user, then MLS deformation method, Eqs. (2.11), (2.12), and (2.13) as described in section 2.4.1.3, is applied to find the optimal transformation for all vertices of the generic 3D mesh. Figure 2.13 illustrates 2D projections of silhouette of the generic 3D model of the femur (green color) on patient 2D bi-planar radiographs. In Figure 2.13, (A) shows initial position of the MLS handles (blue dots) after manual initialization of the generic 3D model of the femur on the patient's radiographs, and (B) shows the manual adjustment of the MLS handles to fit green contours to the femur edges in the patient's radiograph. In Figure 2.13, (C) illustrates 2D projection of the 3D silhouette of the personalized 3D model of the femur on the patient's radiographs.

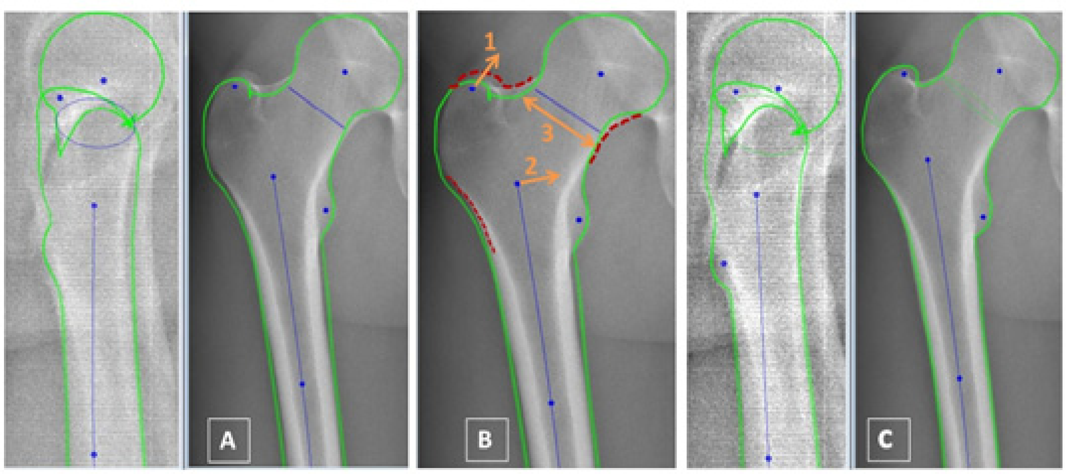


Figure 2.13 2D projections of silhouette of the 3D femur model on patient 2D radiographs
Taken from Chaibi et al. (2012)

In (Cresson et al., 2010), three MLS handles of the 3D femur including the femoral head and two condyles are manually initialized and adjusted by an operator. Then, an ICP-based similarity registration between 3D contours of the generic model and 3D edges of the femur is applied to estimate 3D positions of the MLS handles. Ultimately, to deform the 3D model of the femur, the new positions of the vertices on the 3D mesh of the femur are computed using the interpolation-based MLS deformation method. In 3D femur deformation, MLS deformation drives a constrained and as-rigid-as-possible local 3D shape deformation on a small set of 3D handles and avoids undesirable distortion in the entire generic 3D model, which is suitable for a real-time 3D bone reconstruction (Chaibi et al., 2012 ; Cresson et al., 2008, 2010). In (Chaibi et al., 2012), a small set of 17 handles are manually displaced over the whole femur by an operator, and the reported mean reconstruction time for both lower limbs is 10 minutes, with CPU computation and the mean of the 3D reconstruction accuracy is 1.2 mm. However, the 3D femur reconstruction in (Chaibi et al., 2012) suffers from operator intervention, a lack of reproducibility, and high time consumption. In contrast, (Cresson et al., 2010) proposes a contour-based iterative optimization of the 17 MLS handles for 3D femur reconstruction and the mean of the 3D femur reconstruction accuracy is 1.0 mm. However, (Cresson et al., 2010) suffers from manual initialization. In (Yu et al., 2016), the interpolation-based free-form deformation (FFD) is used for control points-based 3D shape deformation and the mean of the 3D proximal femur reconstruction accuracy is 1.3 mm. In contrast to MLS deformation, which allows physicians to easily adjust possible errors, a large set of 3D control points (88) are uniformly distributed over the entire 3D volumetric template, making any further adjustments almost impossible. This makes it hard to deploy the automatic method of (Yu et al., 2016) in clinical routine. Moreover, in rigid bone deformation, FFD is likely to produce undesirable distortion and needs strong regularization (Yu et al., 2016). Of note, MLS deformation could avoid such undesirable distortion (Cuno & Esperan, 2007).

2.4.4 Deep learning-based 3D/2D non-rigid registration

Over the last decade, state-of-the-art methods propose deep learning-based methods for non-rigid image registration. Unlike an iterative optimizing of (dis)similarity metric (Baka et al., 2012 ; Youn et al., 2017 ; Yu et al., 2016), deep leaning-based methods create a direct map between input image appearances and transformation parameters (Sokooti et al., 2017). Deep convolutional neural networks (CNN) present a high performance in learning non-linearity of mapping function between input images and transformation parameters. (Haskins et al., 2020) widely review deep learning-based 3D/2D non-rigid registration approaches.

Figure 2.14 illustrates an instance architecture used to predict the deformation field for deformable medical image registration.

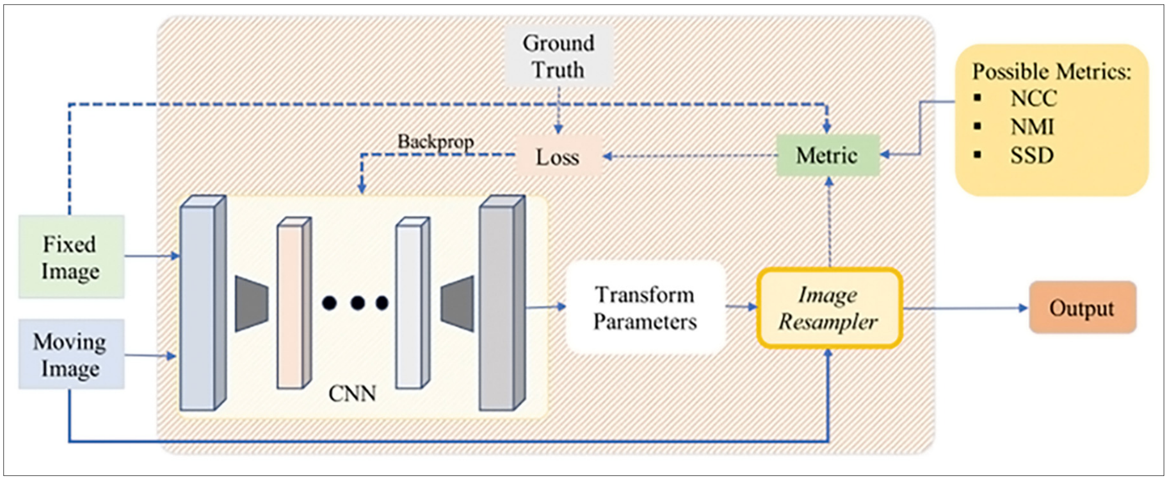


Figure 2.14 CNNs-based regression model with dual supervision in which loss function integrates both ground truth and similarity metric
Taken from Haskins et al. (2020)

In a fully automatic SSM-based fast 3D spine reconstruction method based on the EOS® 2D bi-planar radiographs, (Aubert et al., 2019) proposes a CNN-based regression model to estimate 2D displacement fields of stereo corresponding landmarks to deform 3D statistical shape model of vertebral. The two-channel CNN-based regression model creates a direct map from input pair of (frontal + lateral) local patches, centered at 2D projection of stereo corresponding landmarks, to 3D displacement on frontal and lateral patches (Figure 2.15). Top channel extract contextual information around the vertebra and bottom channel extract local intensity information inside the local patches. The new 3D position of the VBC is used to deform the spine SSM model. The mean error of VBC location is 1.6 mm. Using intensity features allows training CNN-based regression models to deform SSM. Using CNN-based regression models helps to avoid local optimal of conventional optimization of intensity-based non-convex objective function and to obtain acceptable 3D reconstruction accuracy. Using CNN leads to achieve a fast 3D spine reconstruction time, which is less than one minute. However, SSM-based 3D bone reconstruction method needs a training population, finding corresponding points, and good initialization of points to avoid local optimal in optimizing 3D deformable shape parameters. Hence, using the generic model instead of SSM helps to get rid of 3D shape initialization and enable us to use MLS deformation method which provides more flexibility in 3D bone deformation without undesirable distortion.

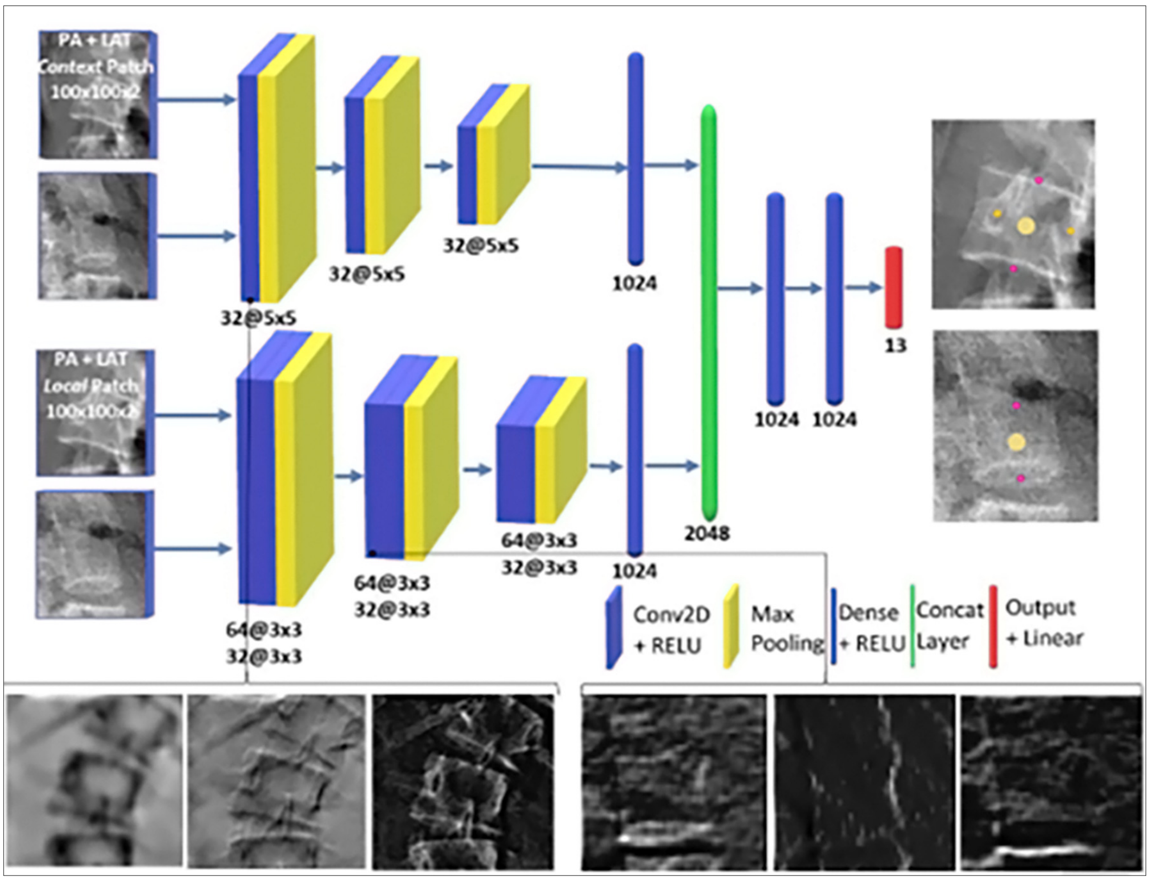


Figure 2.15 CNN-based regression model to estimate 2D displacements fields of stereo corresponding landmark (VBC) on frontal and lateral patches (yellow points)
Taken from Aubert et al. (2019)

As another example, unlike SSM-based 3D reconstruction methods , in an automatic 3D knee bone reconstruction from 2D bi-planar radiographs (Kasten, Doktofsky, & Kovler, 2020), supervised and non-supervised CNN-based models are used for 3D segmentation of the knee bone. Instead of using SSM of the knee bone which need accurate initialization points, the CNN learns to directly reconstruct 3D models of the knee from input DRRs, rendered from CT-scans, (Figure 2.16). The 3D reconstruction process for the proximal femur (red bone in Figure 2.16) takes 45 seconds. Hence, the CNN helps to obtain a fast 3D reconstruction. The mean of Dice and Chamfer errors for distal femur (red bone in Figure 2.16) is 0.943% and 1.075 mm, respectively. However, CT-scan acquisitions provide high level of radiation doses and costs. Using CNN in 3D femur reconstruction leads to obtain comparable and acceptable 3D reconstruction accuracy in a time efficient manner. However, to train the CNN, this method requires CT-scan acquisition and segmentation which provides high levels of radiation dose and with patients in reclining position. Moreover, in the shape deformation analysis of the femur, 3D femur model reconstruction based on the 2D bi-planar radiographs, which are acquired from patient in weight-bearing standing position is very impotent. In addition, the method in (Kasten et al., 2020) does not use the prior 3D model which is not appropriate for our problem since we need a generic 3D model of the femur to personalize and then measure the clinical 3D geometrical parameters.

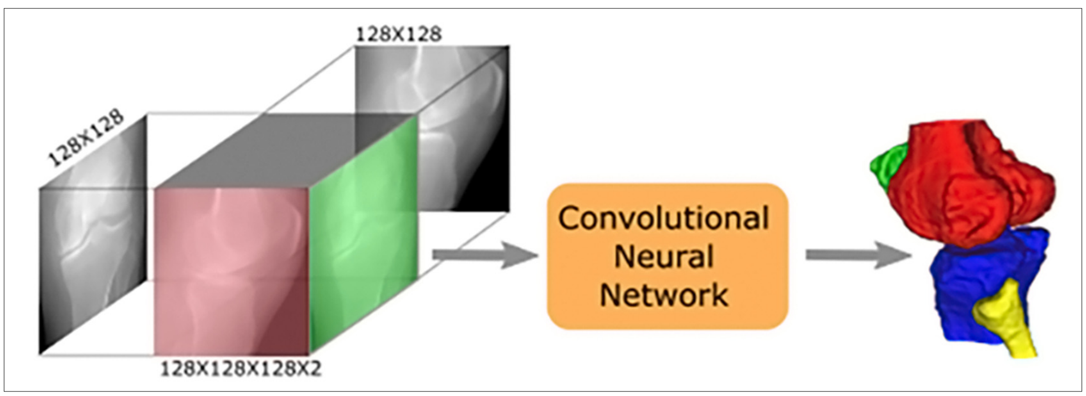


Figure 2.16 Global workflow of CNN-based 3D knee bone reconstruction
Taken from Kasten et al. (2020)

2.4.5 Conclusions on 3D model reconstruction

In 2D radiograph-based 3D femur reconstruction, state-of-the-art methods are improved from semi-automatic (Baka et al., 2011 ; Chaibi et al., 2012 ; Youn et al., 2017) to automatic (Yu et al., 2016). In 3D model reconstruction, a known prior 3D model such as a CT-scan or MRI-based 3D model (Abebe et al., 2011 ; Fang et al., 2020), a statistical shape model (SSM) (Baka et al., 2011, 2012 ; Youn et al., 2017 ; Yu et al., 2017), or a generic template model (Chaibi et al., 2012 ; Cresson et al., 2010 ; Khameneh et al., 2021 ; Yu et al., 2016) is registered into the patient's calibrated 2D bi-planar radiographs. The choice of the prior deformable 3D model drives the 3D bone reconstruction process. SSM-based 3D bone reconstruction methods need a training population, finding corresponding points, and good initialization of points to avoid local optimal in optimizing 3D deformable shape parameters. Unlike PCA-based 3D reconstructions, which require a training process and collecting a CT-scan-based data set (Baka et al., 2011, 2012 ; Fleute & Lavallée, 1999 ; Yu et al., 2017, 2015), MLS-based methods (Chaibi et al., 2012 ; Cresson et al., 2010) construct personalized 3D bone from a single CT-scan-based reconstructed generic 3D model of the target structure without statistical knowledge of the population. In contrast to the PCA-based methods, which globally constrain 3D shape deformation to the training population, an MLS-based method provides a flexible and local 3D shape and scale deformation (Chaibi et al., 2012 ; Cresson et al., 2010). MLS deformation drives a constrained and as-rigid-as-possible local 3D shape deformation on a small set of 3D handles and avoids undesirable distortion in the entire generic 3D model, which is suitable for a real-time 3D bone reconstruction (Chaibi et al., 2012 ; Cresson et al., 2010). Hence, using the generic model instead of SSM helps to get rid of 3D shape initialization and enable us to use MLS deformation method which provides more flexibility in 3D bone deformation without undesirable distortion. In addition, in free-form deformation (FFD) (Yu et al., 2016), a large set of 3D control points (88) are uniformly distributed over the entire 3D volumetric template, which makes any further adjustments almost impossible. Hence, this makes it hard to deploy the automatic FFD-based method in (Yu et al., 2016) in clinical routine. In contrast, MLS deformation allows physicians to easily adjust possible errors by moving a small set of 17 handles and is already used in clinical routines as it provides a user-friendly 3D model adjustment. Moreover, in rigid bone deformation, FFD is likely to undesirable distortion and needs strong regularization (Yu et al., 2016). On the other hand, MLS deformation avoids undesirable distortion (Cuno & Esperan, 2007). MLS-based 3D femur reconstruction method, which is integrated in the commercial SterEOS software (Chaibi et al., 2012 ; Cresson et al., 2010), is an user-friendly tool and is currently used in clinical routine. This tool is semi-automatic, and nonetheless has certain limitations, such as the operator's skill dependency, reproducibility, and time consumption. (Cresson et al., 2010) attempted to automate the MLS deformation, integrated in SterEOS, using automatic ICP-based 3D pose estimation. This approach suffers from manual initialization and correspondence finding between 3D silhouette of the generic 3D femur and 3D contours of the femur corresponding to 2D contours in frontal and lateral radiographs. However, this semi-automatic 3D femur reconstruction approach can be automated in order to remove its current limitations.

In automatic and fast 3D bone reconstruction applications, recent developments of CNN-based methods are inspiring and they show promising results on assigned tasks without user interventions (Aubert et al., 2019; Kasten et al., 2020). Over the last decade, in many 2D radiograph-based 3D model reconstruction applications, deep learning (DL)-based methods have been useful tools for automatic and fast non-rigid registration (Aubert et al., 2019; Haskins et al., 2020). For instance, in fully automatic 2D bi-planar radiographs-based 3D spine reconstruction (Aubert et al., 2019), a CNN-based regression model successfully estimates 3D displacements of corresponding stereo landmarks and the vertebral body center (VBC) from input (frontal+lateral) local patches. Therefore, in 3D/2D non-rigid registration of the femur, using CNN-based regression models is a very interesting and useful tool in order to remove operator interventions and automatically estimate 3D displacements of the 17 MLS handles. In addition, the proposed CNN architecture by (Miao et al., 2016 ; Zheng et al., 2018) is so inspiring and applicable for 3D displacement estimation of MLS handles, since this CNN structure is able to successfully estimate 3D transformation parameters from input local intensity residuals computed in extracted local patches to rigidly register 3D object to 2D radiographs. Hence, in 3D/2D non-rigid registration of the generic femur into 2D radiographs, the same CNN architecture can be used to estimate 3D displacements of MLS handles.

2.5 Evaluation of 3D/2D registration in 3D bone reconstruction

In 3D bone reconstruction applications, evaluation of the 3D/2D registration is an integrated part of the state-of-the-art methods (Goswami & Kr., 2015 ; Hosseinian & Arefi, 2015 ; Markelj, Tomaževič, Likar, et al., 2012 ; Reyneke et al., 2019). Evaluation of the 3D/2D registration leads to determine the performance, limitations, and the potential of using a proposed method in clinical applications (Markelj, Tomaževič, Likar, et al., 2012). In 3D bone reconstruction, the state-of-the-art methods evaluate the accuracy of the 3D pose estimation and 3D/2D non-rigid registration by providing prerequisites such as:

- Gold standard 3D model as a ground truth (Markelj, Tomaževič, Likar, et al., 2012);
- Fuzzy gold standard 3D model (Jannin et al., 2002);
- Evaluation metrics (Markelj, Tomaževič, Likar, et al., 2012);
- 3D pose parameters accuracy including translation, rotation, and isotropic scaling (Baka et al., 2011 ; Miao et al., 2016);
- Clinical 3D geometric parameters accuracy (Chaibi et al., 2012 ; Markelj, Tomaževič, Likar, et al., 2012).

2.5.1 Gold standard 3D model as ground truth

In 3D bone reconstruction applications, the gold standard 3D model, as a ground truth, is the CT-scans-based 3D reconstructed model. For instance, (Baka et al., 2011 ; Chaibi et al., 2012 ; Cresson et al., 2010 ; Youn et al., 2017 ; Yu et al., 2016) use the corresponding CT-scans-based 3D model as ground truth to evaluate 3D shape accuracy of the reconstructed 3D bone model. (Baka et al., 2011, 2012) use CT-scans-based 3D model as ground truth to evaluate 3D pose accuracy.

2.5.2 Fuzzy gold standard 3D model

To evaluate the 3D pose and 3D shape of the reconstructed 3D bone model, stat-of-the-art methods might use fuzzy gold standard 3D model. For example, an expert constructs personalized fuzzy gold standard 3D model of the target bone structure using semi-automatic 3D bone reconstruction method (Chaibi et al., 2012) integrated in commercial software tool (SterEOS).

2.5.3 Evaluation metrics

In 3D bone reconstruction applications, to evaluate 3D pose parameters, 3D shape, and clinical 3D parameters, state-of-the-art methods use different evaluation metrics. For instance, to evaluate accuracy of the estimated 3D pose parameters, (Baka et al., 2011, 2012) compute absolute pose difference between ground truth and estimated 3D pose parameters. In 3D/2D registration of distal femur, (Baka et al., 2011) computes absolute pose differences metric to evaluate 3D translations, 3D rotations, and isotropic scaling estimation errors in comparison to ground truth 3D pose parameters. Absolute pose error is measured by absolute subtraction between the estimated and ground truth 3D pose parameters (Baka et al., 2011). Then, the Root Mean Square Error (RMSE) (Hyndman & Koehler, 2006) on validation set is computed for each 3D pose parameter (Baka et al., 2011, 2012). The RMSE is a standard way to measure the error of a model in estimating quantitative data, and is computed as, Eq. (2.17):

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y_i} - y_i)^2}{n}} \tag{2.17}$$

Where n is the number of observations, $\hat{y_i}$ is the ground truth 3D pose parameter value, and y_i is the estimated 3D pose parameter value. Another metric to evaluate 3D pose estimation is the success rate, which shows in how percentage the registration error is less than pre-defined threshold (Miao et al., 2016). In a 3D/2D rigid registration of a rigid object, (Miao et al., 2016) computes Mean target registration error ($mTRE_{proj}$) in (mm), Eq. (2.18), which is calculated at the 8 corners of two bounding boxes surrounded the rigidly registered 3D model, and the gold standard 3D model, Figure 2.17 (Van De Kraats, Penney, Tomaževič, Van Walsum, & Niessen, 2005).

$$mTRE_{proj}(p, T_{reg}, T_{gold}, \hat{n}) = \frac{1}{k} \sum_{i=1}^{k} \left\| \left(T_{reg} p_i - T_{gold} p_i\right) \cdot \hat{n} \right\| \tag{2.18}$$

where $mTRE_{proj}$ is the mean of target registration error in projection direction, k is the 3D points number which is 8, i shows the 3D point index, T_{reg} determines the 3D transformation parameter estimated by 3D/2D registration algorithm, T_{gold} is the gold standard 3D transformation parameters, and $\hat{n}$ shows the normal of the projection plane (Van De Kraats et al., 2005).

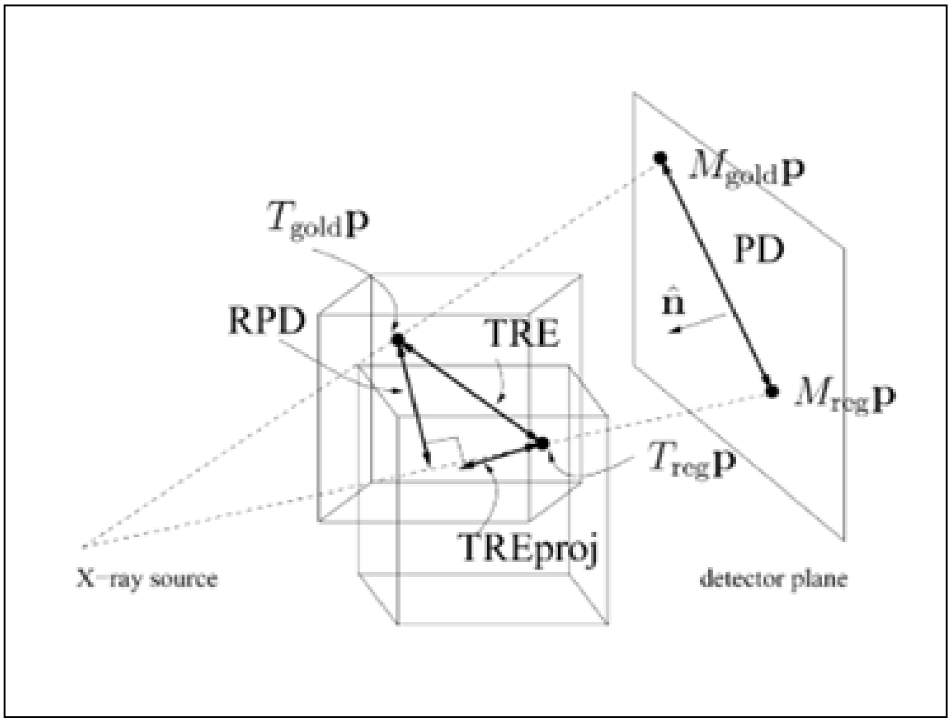


Figure 2.17 Depiction of computation of mTRE in projection direction
Taken from Van De Kraats et al. (2005)

In addition, (Miao et al., 2016) measures capture range of 3D pose estimation in the 3D/2D rigid registration. The capture range evaluates the range starting position for 3D translation and 3D rotation in which the registration is successful (Van De Kraats et al., 2005).

To evaluate the 3D bone shape reconstruction accuracy, (Baka et al., 2011, 2012 ; Chaibi et al., 2012 ; Cresson et al., 2010 ; Youn et al., 2017 ; Yu et al., 2016) computes the point-to-surface (P2S) distance between the ground truth and the reconstructed 3D bone model. Then, the Root Mean Square (RMS) (Baka et al., 2011, 2012) and Standard Deviation (SD) of the P2S errors are computed over the evaluation subjects. In (Youn et al., 2017 ; Yu et al., 2015), the Hausdorff distance error, Eq. (2.19), (Rote, 1991) is computed between the ground truth and the reconstructed 3D femur and proximal femur model, respectively. The Hausdorff distance measures how far two 3D point sets are far from each other, as Eq. (2.19), (Figure 2.18) :

$$d_H(X,Y) = \max\left\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\right\} \tag{2.19}$$

where X and Y are two 3D point sets, sup represents supremum or maximum (Figure 2.18). Afterwards, in (Youn et al., 2017 ; Yu et al., 2015), the (RMS $\pm$SD) and mean of the Hausdorff distances are calculated on the evaluation set as 3D reconstruction accuracy, respectively.

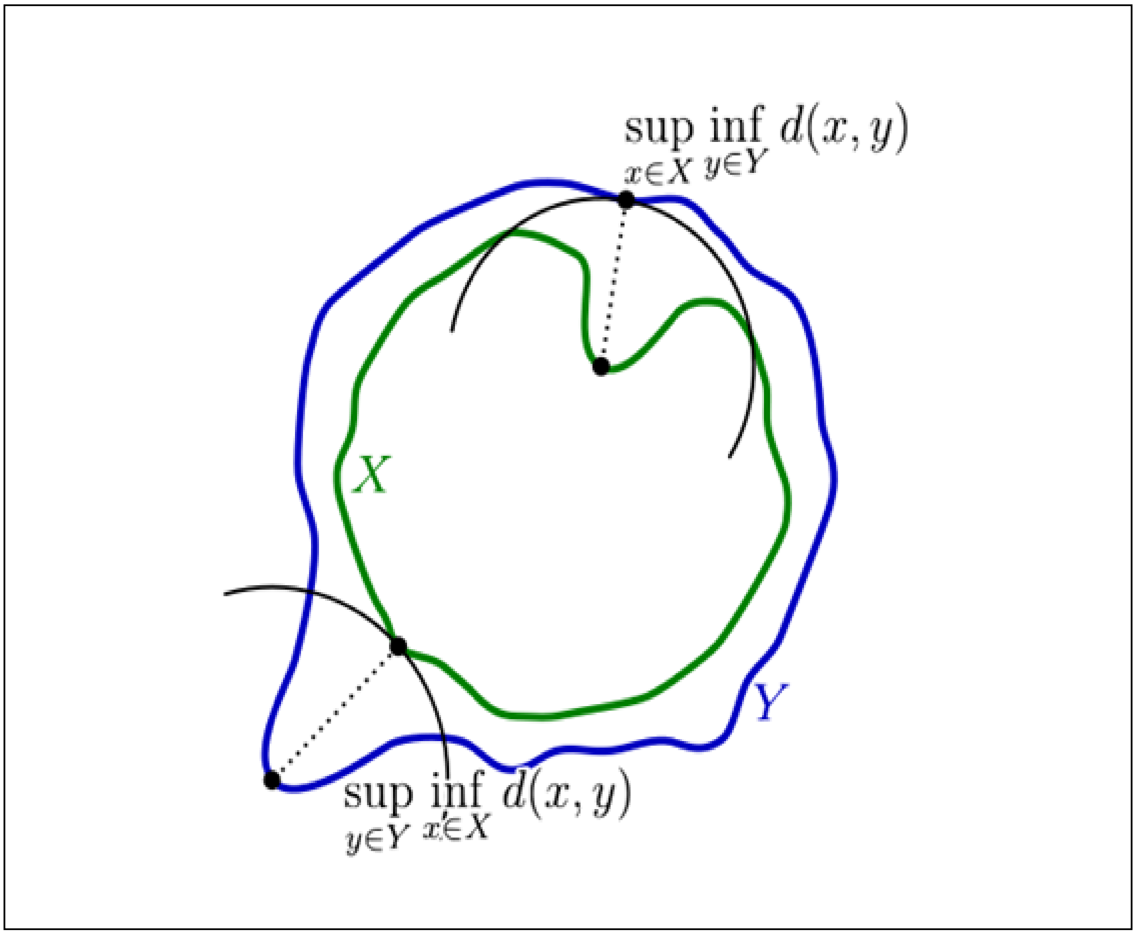


Figure 2.18 Hausdorff distance computation between green (X) and blue (Y) curve
Taken from https://commons.wikimedia.org/wiki/File:Hausdorff_distance_sample.svg

In (Yu et al., 2016, 2017), the accuracy of the reconstructed 3D model of the proximal femur is validated by computing 3D average distance to measure average distance between vertices of the surface of ground truth volume to surface extracted from reconstructed volume. Average surface distance (ASD) (Yu et al., 2016) of 3D cortical bone region is calculated between inner cortical bone surfaces of two binary masks of the manually segmented cortical bone (Yu et al., 2016). A 3D Dice coefficient (DC) (Dice, 1945), Eq. (2.20), is computed to measure the overlap between ground truth CT data binary segmentation, L_1, and binary segmentation of the reconstructed volume, L_2, (Yu et al., 2016). Binary segmentation is done manually (Yu et al., 2016).

$$DC = 2|L_1 \cap L_2|/(|L_1| + |L_2|) \quad (2.20)$$

A 3D cortical bone Dice coefficient (CBRDC) (Yu et al., 2016) is calculated to measure the result of control-based 2D/3D non-rigid registration of cortical bone based on DC metric (Yu et al., 2016). By using deformation field, the binary mask of the cortical bone region of volumetric template is warped to get the reconstructed cortical bone. Then, this reconstructed cortical bone is compared with the reference cortical bone, extracted from ground truth CT, using DC metric. In (Jianhua Yao & Taylor, 2003), the accuracy of the reconstructed 3D volume of the pelvic is evaluated by computing volume overlap percentage error, Eq. (2.21), to compare reconstructed volumetric model and ground truth volumetric model. They scan the volumetric model in X, Y, and Z axis to obtain a set of isotropic voxels inside the volumetric model. Then, the rate of overlapping voxels to the total number of voxels is computed as volume overlap percentage. V_g is the set of voxels in the ground truth model and V_s is the set of voxels in the reconstructed model, and $\|\cdot\|$ represents the seize of the set (Jianhua Yao & Taylor, 2003).

$$Overlap = \|V_g I\ V_s\|/\|V_g\| \times 100\% \quad (2.21)$$

In addition, the state-of-the-art methods reports the computation time of 3D bone reconstruction (Aubert et al., 2019 ; Chaibi et al., 2012 ; Cresson et al., 2010 ; Laporte et al., 2003 ; Yu et al., 2017, 2015).

To evaluate clinical 3D geometrical parameters, in (Aubert et al., 2019 ; Chaibi et al., 2012), after 3D model reconstruction, the accuracy of clinical geometric 3D parameter measurements based on the reconstructed 3D model are evaluated in comparison with the ground truth model. In the lower limb 3D bones reconstruction, the most important 3D clinical measurements include (Guenoun et al., 2012):

- femur length, tibia length;
- lower limbs length;
- hip knee center-femoral shaft angle (HKS) angle;

- hip-knee-ankle angle (HKA) angle;
- femoral head diameter;
- femoral mechanical angle (FMA);
- femoral length (FL);
- femoral rotation (FT);
- tibial torsion.

(Chaibi et al., 2012) computes mean and standard deviations of clinical geometric 3D parameter measurements' errors over validation cases for HKS (º), FMA (º), FT (º), and FL (mm). (Aubert et al., 2019) computes scoliosis parameters of 3D spine model as clinical measurements along with mean, standard deviations, and mean absolute errors over validation cases.

CHAPITRE 3

RESEARCH PROBLEMATICS, OBJECTIVES, AND PROPOSED METHODOLOGY

The literature review on 3D femur reconstruction reveals the depth of this problem, which still remains some important limitations despite the many works carried out on this subject for the past two decades. Semi-automatic methods, such as the one employed by the EOS® 3D model reconstruction system (Chaibi et al., 2012), require the manual intervention of an operator for pose initialization and shape and scale adjustment of the 3D model to images (Baka et al., 2011, 2012 ; Chaibi et al., 2012 ; Fleute & Lavallée, 1999 ; Laporte et al., 2003 ; Mahfouz, Badawi, Fatah, Kuhn, & Merkl, 2006 ; Mitton et al., 2000 ; Youn et al., 2017). These manual interventions impact time efficiency and reproducibility of the different approaches (Chaibi et al., 2012). Recently, efforts have been deployed to remove any intervention by the operator (Yu et al., 2016) and improve the time efficiency (Aubert et al., 2019) of the registration approaches. Although full automation is highly desired, there is always the potential for errors in the 3D reconstruction process that need to be manually corrected to facilitate deployment in clinical practice. From a practical point of view, the absence of an automatic method with effective and easy correction tool to remain robust in the face of 3D femur reconstruction difficulties is one of the major limitations observed in the literature (Baka et al., 2011, 2012 ; Chaibi et al., 2012 ; Fleute & Lavallée, 1999 ; Laporte et al., 2003 ; Mahfouz et al., 2006 ; Mitton et al., 2000 ; Youn et al., 2017 ; Yu et al., 2017, 2015). Several of the described approaches use the SSM for 3D femur reconstruction, however, it is difficult to collect and cover the morphological variability encountered in clinical needs. Hence, these methods fail to produce satisfactory shape deformation, and the operator needs to correct the reconstructed 3D femur, often with correction tools. The MLS deformation drives a constrained and as-rigid-as-possible local 3D shape deformation on a small set of 3D handles and avoids undesirable distortion in the entire generic 3D femur, which is suitable for a real-time 3D bone reconstruction (Chaibi et al., 2012 ; Cresson et al., 2010). The MLS deformation method allows physicians to easily adjust possible errors, which makes it easy to deploy an automatic method in clinical routine. Presently, a semi-automatic MLS-based 3D femur reconstruction method is integrated in the commercial SterEOS software tool (Chaibi et al., 2012) and is currently used in clinical routine. Although this semi-automatic approach provides physicians with an easy and user-friendly way to manually carry out adjustments to correct potential 3D reconstruction errors, it has nonetheless certain limitations, such as dependency on the operator's skill to manually move the MLS handles, limited reproducibility, and high time consumption. In 3D femur reconstruction, (Cresson et al., 2010) attempt to automate MLS deformation by proposing an ICP-based method to automatically move MLS handles, however, it suffers from manual initialization. This semi-automatic 3D femur reconstruction approach could be automated in order to remove its current limitations.

The main objective of thesis is to present a fully automatic 3D femur reconstruction method to fit a generic 3D femur model into the patient's EOS® 2D bi-planar radiographs and accurately assess clinical 3D geometric parameters in a time-efficient manner. The sub-objective of thesis is to meet the clinical need for an automatic method which provides physicians, if required, with the means to quickly and easily adjust potential errors of the reconstructed 3D femur, to facilitate deployment in clinical practice.

To achieve these objectives, this thesis presents a fully automatic 3D femur reconstruction method with the ability to easily correct the reconstructed 3D model's potential errors, if required, via a small number of MLS handles. The proposed automatic 3D/2D registration framework combines deep Convolutional Neural Networks (CNN) cascade-based registration models and the Moving Least Square (MLS) deformation to fit the generic 3D femur model into the patient's EOS® 2D bi-planar radiographs. The main contributions of this thesis include:

1) Automatic 3D pose and isotropic scale estimation of the femur;
2) Merging the estimation of CNN-based handles' 3D displacements and scale ratios with MLS handles' deformation to automate the 3D/2D non-rigid registration process.

This project includes two main stages consisting of:

1) 3D pose estimation (3D/2D similarity registration);
2) 3D/2D non-rigid registration.

Chapter 4 describes in detail the proposed methodology of the thesis.

CHAPITRE 4

AUTOMATIC 3D FEMUR RECONSTRUCTION

4.1 Introduction

In a personalized 3D femur reconstruction, one of the key steps is geometrical relationship estimation between an anatomical template 3D model of the bone and the two radiographs of the patient. This crucial process is generally done by firstly estimating a rigid or similarity transformation, and then applying a deformation algorithm to obtain the final reconstruction. Figure 4.1 illustrates global workflow of the proposed fully automatic EOS® 2D bi-planar radiographs-based lower limb 3D bone reconstruction methodology. The workflow consists of two main stages to solve automatic 3D femur reconstruction problem. In the first stage, 3D/2D similarity registration deals with 3D femur pose and isotropic scale estimation U_{Pose} to register the generic 3D model into the patient's EOS® 2D bi-planar radiographs. In second stage, 3D/2D non-rigid registration estimates local 3D shape deformation U_{Shape} and local 3D scale deformation U_{Scale} .

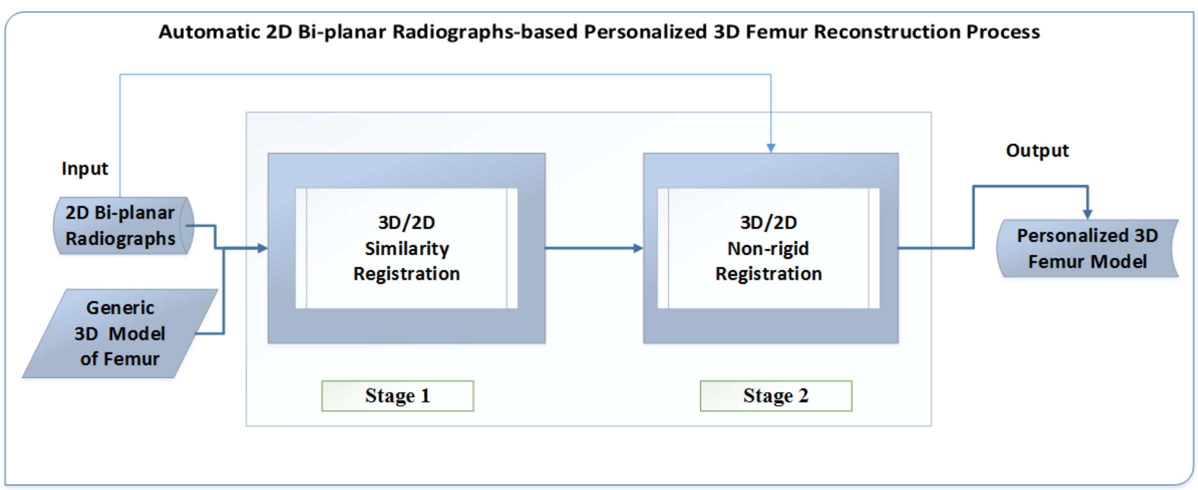


Figure 4.1 Overview of the proposed workflow for personalized 3D femur reconstruction from the EOS® 2D bi-planar radiographs

The 3D pose and isotropic scale, local 3D shape, and local 3D scale ratio are sequentially optimized to automatically reconstruct a personalized geometric 3D model of the target femur 3D bone structure.

4.2 Automatic Personalized 3D Femur Reconstruction

In 2D bi-planar radiographs-based personalized 3D femur reconstruction process, Eq. (4.1) defines the objective function to automatically fit the generic 3D model of the femur, (M_g), into (F), which is the patient's target femur 3D bone. The proposed CNN cascade-based 3D/2D registration automatically fits (M_g) into (F), using 2D bi-planar radiographs, as Eq. (4.1):

$$(F) = U_{Scale} \circ \left(U_{Shape} \circ \left(U_{Pose} \circ \ (M_g),\right)\right) \tag{4.1}$$

The CNN cascade-based 3D/2D registration framework (Figure 4.2) consecutively uses three CNN-based regression models, CNN_{Pose}, CNN_{Shape}, and CNN_{Scale} to fit (M_g) into (F) via two main stages:

1) 3D/2D similarity registration (U_{Pose}): in a coarse-to-fine 3D/2D similarity registration, a PCA-based registration is used to coarsely initialize the pose of the generic 3D model of the femur. The PCA is applied on CNN-based segmented masks of the detected femur in the EOS® 2D bi-planar radiographs to recover 3D similarity transformation parameters, including translation, rotation and scale. In the refine registration step, CNN-based regression models are trained to obtain more accurate 3D pose parameters.
2) 3D/2D non-rigid registration, including local 3D shape deformation (U_{Shape}) and local 3D scale deformation (U_{Scale}): to deform the local shape of the registered 3D femur model, CNN-based regression models are trained to find 3D displacements of a small number of 3D handles pre-defined on the femur. Following the computation of the new positions of 3D handles, an MLS deformation is applied to obtain a 3D model better adjusted to the radiographs without any user interactions. To optimize

the local 3D scale corresponding to each handle, CNN-based regression models are trained to estimate the local 3D scale ratios corresponding to the said handles. Then, the MLS method is used to compute 3D rescaling fields and extend the scales to the entire 3D femur.

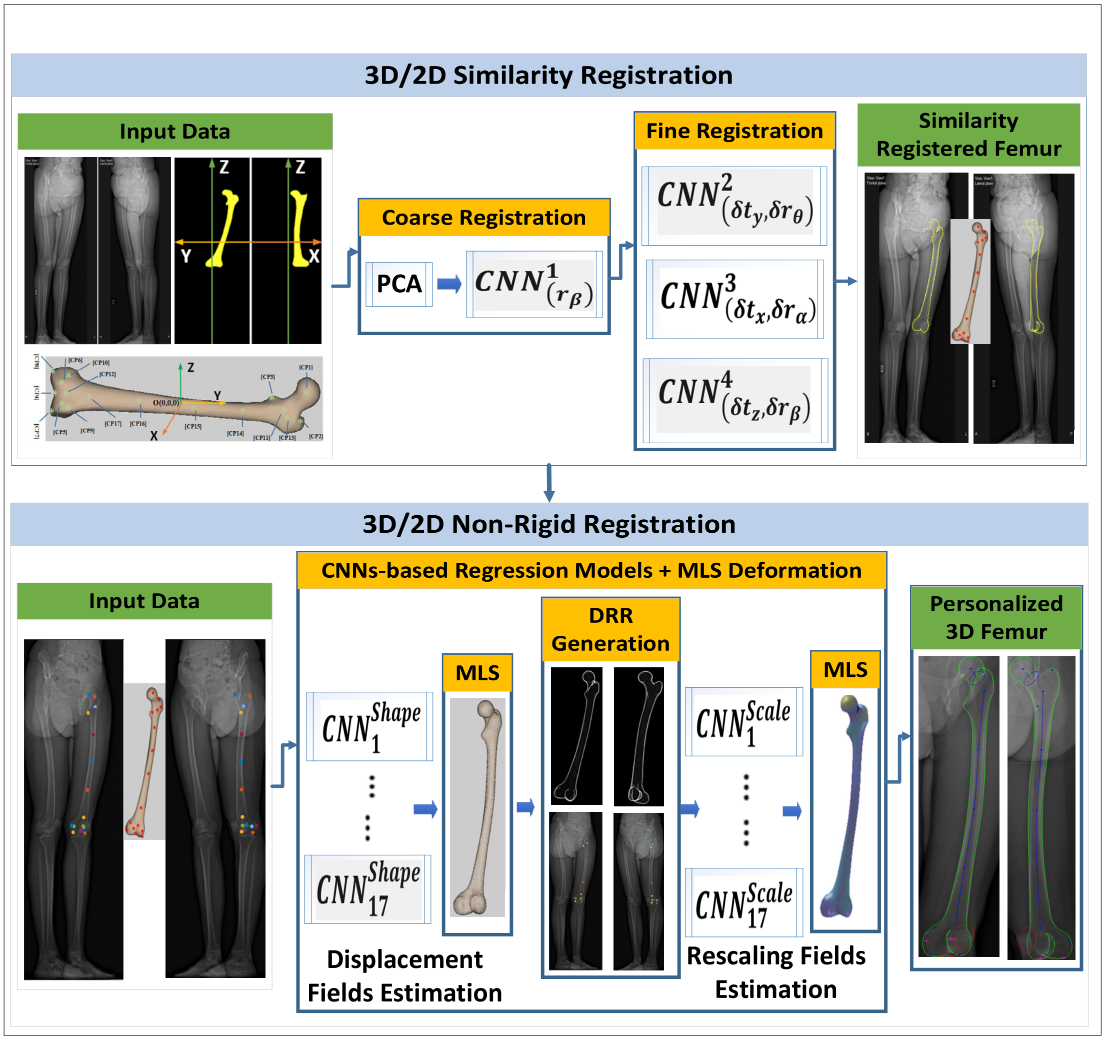


Figure 4.2 Automatic 3D/2D registration framework in two main stages 3D/2D similarity registration (top row), and 3D/2D non-rigid registration (bottom)

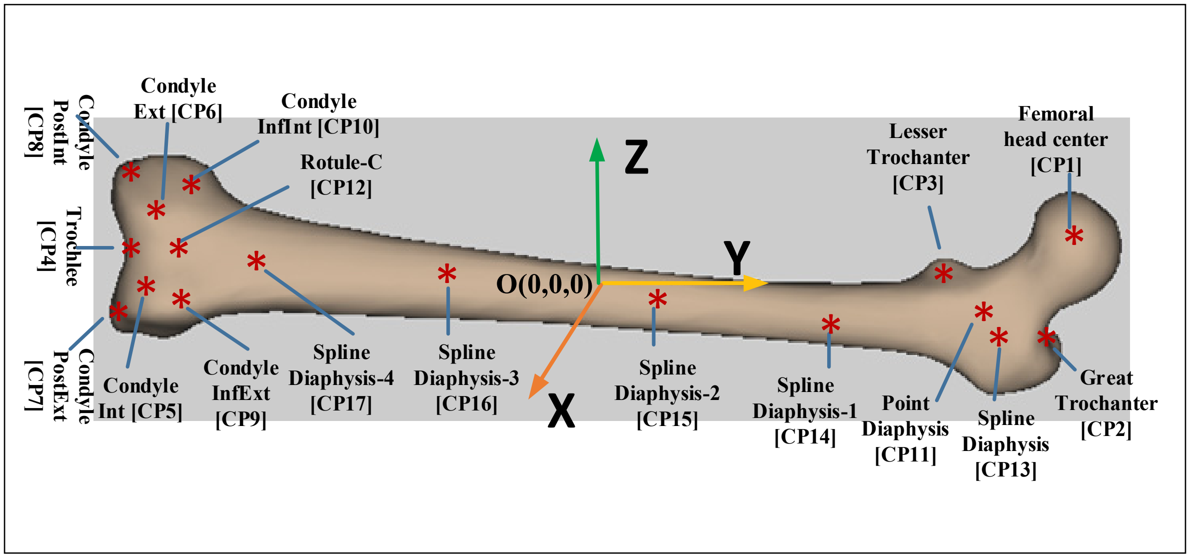


Figure 4.3 Femoral shape description with 17 3D handles at the origin of the 3D referential of the EOS® cabin system

4.2.1 Deformable generic 3D model and 3D coordinate system

To begin, the position of the generic 3D model (M_g) and the patient's bone structure are defined with respect to the origin of the 3D referential of the EOS® cabin system (0,0,0), (Figure 4.3). We use the same generic 3D model $\left(M_g\right)$ constructed by (Chaibi et al., 2012 ; Cresson et al., 2010), which is a computerized tomography scan-based reconstructed 3D surface mesh of the left femur. A set of 17 3D handles recorded as $\boldsymbol{CP} = \{CP_i \in \mathbb{R}^3 | i = 1, \cdots, 17\}$, where i is the index of the 3D handles, are previously defined by (Chaibi et al., 2012) on the generic 3D model $\left(M_g\right)$, (Figure 4.3), to manipulate and control the 3D bone shape deformation. Of the 17 3D handles, CP_1 to CP_{12} are 3D point handles with uniform scales, which are located on the proximal and distal femoral regions. $CP_{13}$to CP_{17} are the spline handles with non-uniform scales, which are located on the femoral diaphysis. The latter handles are more appropriate for depicting the 3D shape deformation of the femoral diaphysis (Chaibi et al., 2012). The 3D mesh $\left(M_g\right)$ is constructed by a set of 3D vertices $\left(V_g\right)$, and each CP_i is surrounded by a corresponding set of vertices called V_i.

4.2.1.1 Moving least square (MLS) deformation of generic 3D model

To deform the generic 3D model, which is controlled by 17 handles, we apply the MLS deformation method (Chaibi et al., 2012 ; Cresson et al., 2010). Given two sets of 3D positions for 17 handles, $\boldsymbol{CP^T}$(source) and $\boldsymbol{CP'}$(reference), the MLS deformation method locally computes the new positions to the corresponding set of the 3D vertices, V_i, of (M_g^T). The Weighted Least Squares (WLS) optimization, Eq. (4.2) provides a locally constrained shape deformation based on a weight function, w_i:

$$\widehat{D}_{V_i} = \underset{\widehat{D}_v}{argmin}\left(\sum_i^{17} w_i \left\| CP'_i - D_{V_i}(CP_i^T) \right\|^2\right), \forall V_i \in \left(V_g^T\right),$$

$$w_i = \frac{1}{d(CP_i^T, V_i)^{2\alpha}} \tag{4.2}$$

where $\widehat{D}_{V_i}$ is an optimal similarity transformation associated with CP_i^T and its corresponding set of vertices V_i and $d(.)$ is the Euclidean distance function between each vertex of the set V_i and CP_i^T. The impact of $d(.)$ on the displacement field of v is controlled by the parameter $\alpha > 0$. In the set of vertices V_i, the vertices that are far from the corresponding MLS handle CP_i^T get low weight in order to have less impact in WLS optimization of the similarity transformation parameters. The optimal similarity transformation $\widehat{D}_{V_i}$ is found using singular value decomposition (SVD) (Arun et al., 1987) by solving a least square regression between the source ($\boldsymbol{CP^T}$) and target ($\boldsymbol{CP'}$) 3D handles. To avoid undesirable distortion and non-uniform scaling, for each handle CP_i^T, $\widehat{D}_{V_i}$ is restricted to similarity transformation as in Eq. (4.3):

$$CP_i^T \rightarrow \widehat{D}_{V_i}(CP_i^T) = sc_{Global} r_i CP_i^T + t_i \tag{4.3}$$

where sc_{Global} is the uniform global scaling and r_i is a local 3D rotation followed by a local 3D translation t_i associated with CP_i^T. For each handle CP_i^T, we transform the corresponding set of vertices V_i with $\widehat{D}_{V_i}$ to obtain a locally constrained and smooth 3D shape deformation on entire $\left(M_g^T\right)$. The deformed 3D model $\left(M_g^{\widehat{D\circ}(T)}\right)$ is used as an input for the next local 3D scale ratio estimation.

4.2.2 Convolutional neural networks (CNN) structure

A CNN cascade-based 3D/2D registration framework comprises 38 CNN-based regression models: four for (U_{Pose}), 17 for (U_{Shape}), and 17 for (U_{Scale}). In the first stage of 3D femur reconstruction, 3D/2D similarity registration (U_{Pose}), four CNN-based regression models are multi-channel CNNs with the same architecture as (Miao et al., 2016), which are trained to estimate 3D transformation parameters' residuals. (Figure 4.4) shows the structure of each multi-channel CNN-based regression model corresponding to (U_{Pose}). The number of channels varies from the coarse to the fine registration step depending on the number of input LIRs. In the coarse registration step, the input of the multi-channel CNN-based regression models is 12 local intensity residuals (LIRs). Hence, we train a twelve-channel CNN-based regression model, $CNN^1_{(r_\beta)}$, corresponding to the input 12 LIRs, where each channel of the CNN corresponds to each LIR. In the fine registration step, the input of multi-channel CNN-based regression models: $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$, is six frontal, six lateral, and 12 bi-planar LIRs, respectively. Each channel of the multi-channel CNN corresponds to each input LIR.

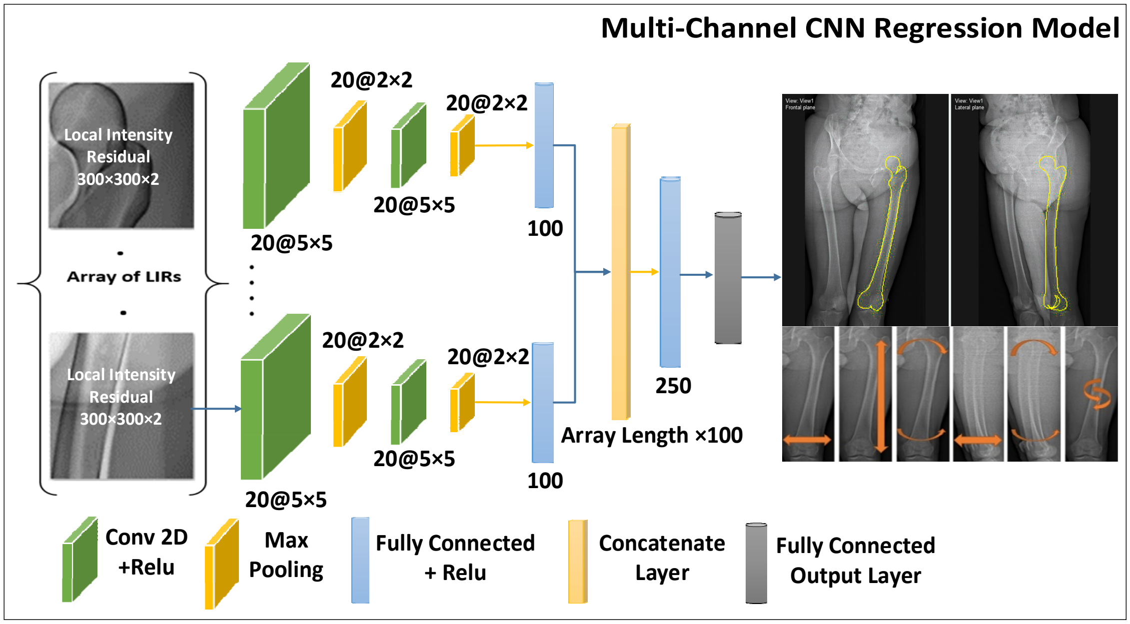


Figure 4.4 Multi-channel CNNs-based regression model's architecture

Then, in the second stage, 3D/2D non-rigid registration with (U_{Shape}) and (U_{Scale}), 34 bi-channel CNN-based regression models, which have the same structure as (Khameneh et al., 2021 ; Miao et al., 2016), are trained to estimate 3D displacements, and 3D scale ratios corresponding to $\boldsymbol{CP}^T$, respectively. (Figure 4.5) shows the structure of a bi-channel CNN-based regression model for (U_{Shape}) corresponding to a MLS handle, CP_i^T. The bi-channel CNN-based regressors of (U_{Scale}) have the same architecture as (U_{Shape}), which is shown in Figure 4.5. In both (U_{Shape}) and (U_{Scale}), each CNN-based regression model has two input local patches (frontal and lateral), Figure 4.5. Since we use both frontal and lateral local patches (bi-planar patches) as input image of the regressors, the CNN-based regressors are bi-channel.

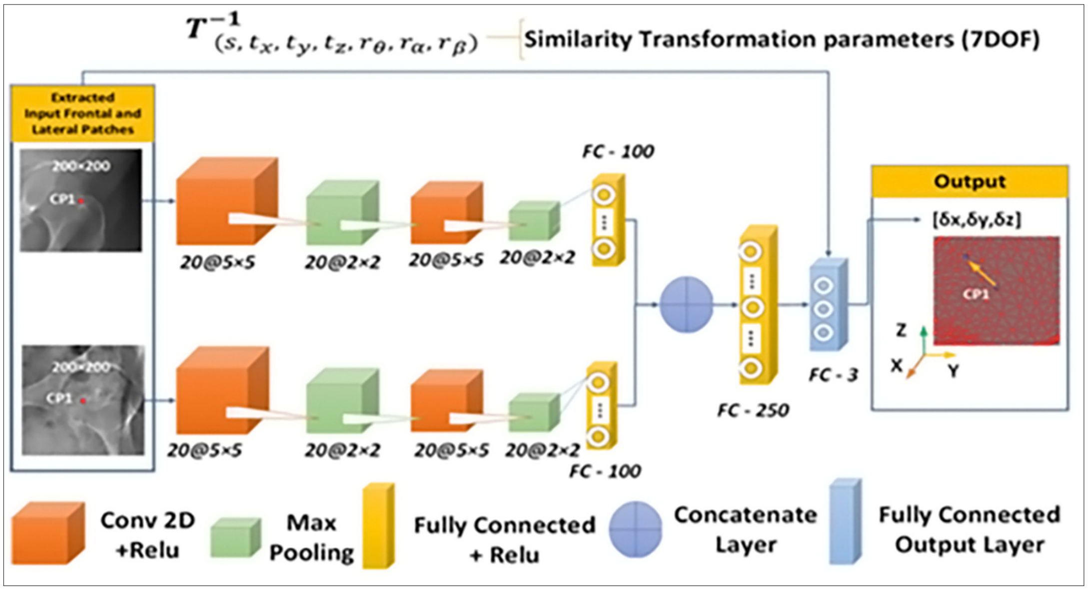


Figure 4.5 Bi-channel CNN-based regression model's architecture

In Figures 4.4 and 4.5, the CNN-based regression model learns a direct map between the input local patches and the desired output. The inputs for (U_{Pose}), (U_{Shape}), and (U_{Scale}) are local intensity residuals (LIRs), local intensity (LI), and local intensity residuals (LIRs), respectively, inside the extracted local patches described in sections 4.3.6, 4.4.2.1, and 4.4.3.1, respectively. The output for (U_{Pose}), (U_{Shape}), and (U_{Scale}) are 3D transformation residuals, 3D displacements, and 3D scale ratios, respectively. In Figures 4.4 and 4.5, each CNN channel is constructed from two convolutional layers followed by drop-out layers, two max pooling layers, and a fully connected layer. The extracted feature maps from each CNN channel are concatenated and passed to another fully connected layer. The next fully connected layer is the output layer. Each CNN-based regression model is trained to minimize the Euclidean loss as an objective function, as in Eq. (4.4):

$$\psi = \frac{1}{n}\sum_{j=1}^{n}\left\|G_j - Y_j\right\|_2^2 \tag{4.4}$$

Where n, is the number of training samples, G_j is the known target, and Y_j is the estimated output for the j^{th} training sample. For (U_{Pose}), (U_{Shape}), and (U_{Scale}), G_j and Y_j are the target and estimated 3D pose parameter residual, 3D displacements, and 3D scale ratios, respectively. In each CNN-based regressor, the Xavier method (Xavier & Bengio, 2010) is used to initialize the weights. In each iteration of training (epoch = 100), the weights are optimized via the Adam optimizer (Kingma & Ba, 2015), the mini-batch size equals to 10, and the learning rate is 0.009.

4.3 Automatic 3D/2D Similarity Registration (U_{Pose})

4.3.1 Introduction

In this thesis, we propose a fully automatic Convolutional Neural Network (CNN)-based 3D/2D similarity registration, U_{Pose}, to estimate 3D femur pose and isotropic scale based on a coarse-to-fine strategy. First, to directly initialize the 3D femur's pose and isotropic scale parameters, an automatic CNN-based semantic segmentation (Agomma, Vazquez, Cresson, & Guise, 2019) is used to segment 2D bi-planar masks of the femur followed by a Principal Component Analysis (PCA)-based rigid registration. Then, CNN-based regression models, Figure 4.4) are trained on pose-invariant local patches' similarities to refine the 3D pose parameters. Figure 4.6 shows seven transformation parameters (7DOF), from left to right, isotropic scaling (s), horizontal translation in frontal plane (t_y), vertical translation in both frontal and lateral plane (t_z), frontal in-plane rotation (r_θ), horizontal translation in lateral plane (t_x), lateral in-plane rotation (r_α), out-of-plane rotation (r_β).

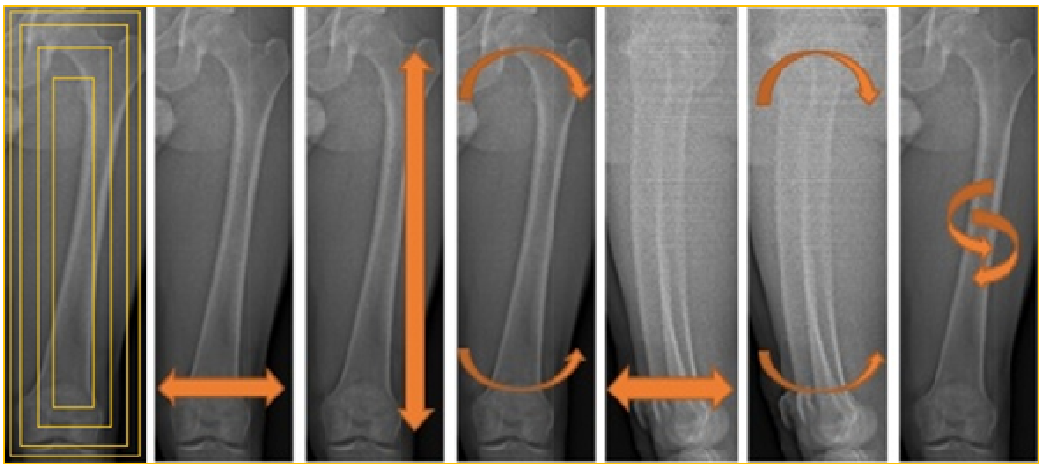


Figure 4.6 Illustration of seven degrees of freedom (7DOF) transformation parameters

In the proposed fully automatic coarse-to-fine CNNs-based 3D/2D similarity registration, U_{Pose}, the homogenous transformation matrix $T(t_x, t_y, t_z, r_\theta, r_\alpha, r_\beta, s)$ aligns the generic 3D model (M_g) into the patient's target femur (F) using 2D bi-planar radiographs via Eq. (4.5):

$$(F) = \delta T \circ \left(T \circ (M_g)\right) \tag{4.5}$$

Where, $T(t_x t_y, t_z, r_\theta, r_\alpha, r_\beta, s)$ is a homogenous transformation matrix comprising 3D translations $(t_x t_y, t_z)$, 3D rotations $(r_\theta, r_\alpha, r_\beta)$, and 3D isotropic scale (s). The coarse registration step initializes T, and then fine registration step estimates δT to refine T, Eq. (4.5).

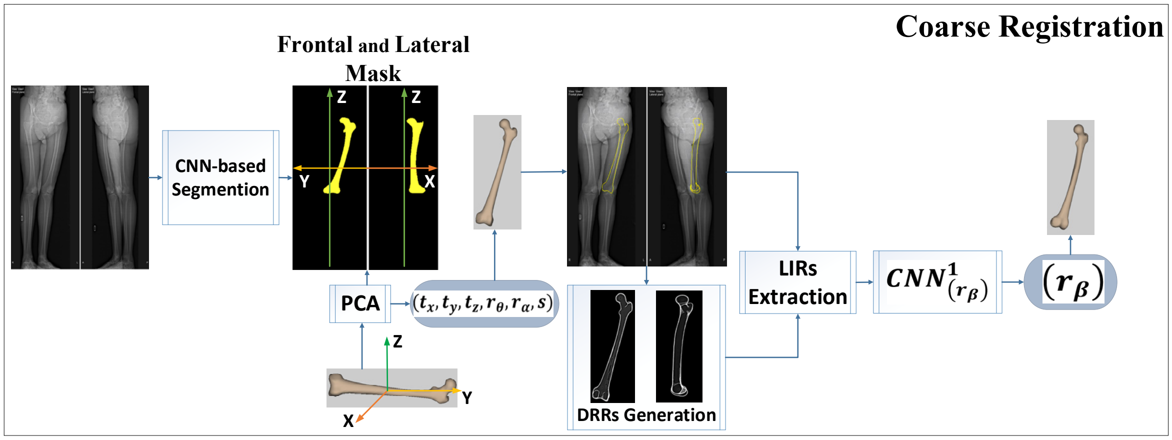


Figure 4.7 The workflow for coarse registration

4.3.2 Coarse registration

The coarse registration directly initializes transformation matrix T, to coarsely register (M_g) into (F), Eq. (4.5). Figure 4.7 shows the workflow for the coarse registration consisting of PCA-based registration to estimate $(t_x t_y, t_z, r_\theta, r_\alpha, s)$ and CNN-based regression to estimate (r_β). The inputs of the coarse registration include: (M_g), 2D bi-planar radiographs of the femur, CNN-based segmented frontal and lateral 2D masks of the femur. The output of the coarse registration is a coarse estimation of seven transformation parameters $(t_x t_y, t_z, r_\theta, r_\alpha, s)$ and a roughly registered generic 3D model to obtain (M_g^T).

4.3.3 PCA-based registration to estimate $(t_x t_y, t_z, r_\theta, r_\alpha, s)$

In PCA-based registration, the input is (M_g) and frontal and lateral segmented 2D masks of the femur, and the output is estimated six transformation parameters $(t_x t_y, t_z, r_\theta, r_\alpha, s)$. Frontal and lateral segmented 2D masks $(Mask_F, Mask_L)$ of the femur (F) are extracted from 2D bi-planar radiographs (I_F, I_L) of the patient via an automatic CNN-based semantic segmentation approach (Agomma et al., 2019). Then, we apply a PCA (Salvi, Matabosch, Fofi, & Forest, 2007) on $(Mask_F, Mask_L)$, to find the 2D inertial centers (C_F^{2D}, C_L^{2D}), and the directions of the principal axis of the mask pixels' distributions $(R_F^{2D},\ R_L^{2D})$. (Figure 4.8) shows application of PCA on $(Mask_F, Mask_L)$ with yellow color, where red dots on $(Mask_F, Mask_L)$ are the 2D inertial centers (C_F^{2D}, C_L^{2D}) and black lines on $(Mask_F, Mask_L)$ show the directions of the principal axis of the mask pixels' distributions (R_F^{2D}, R_L^{2D}).

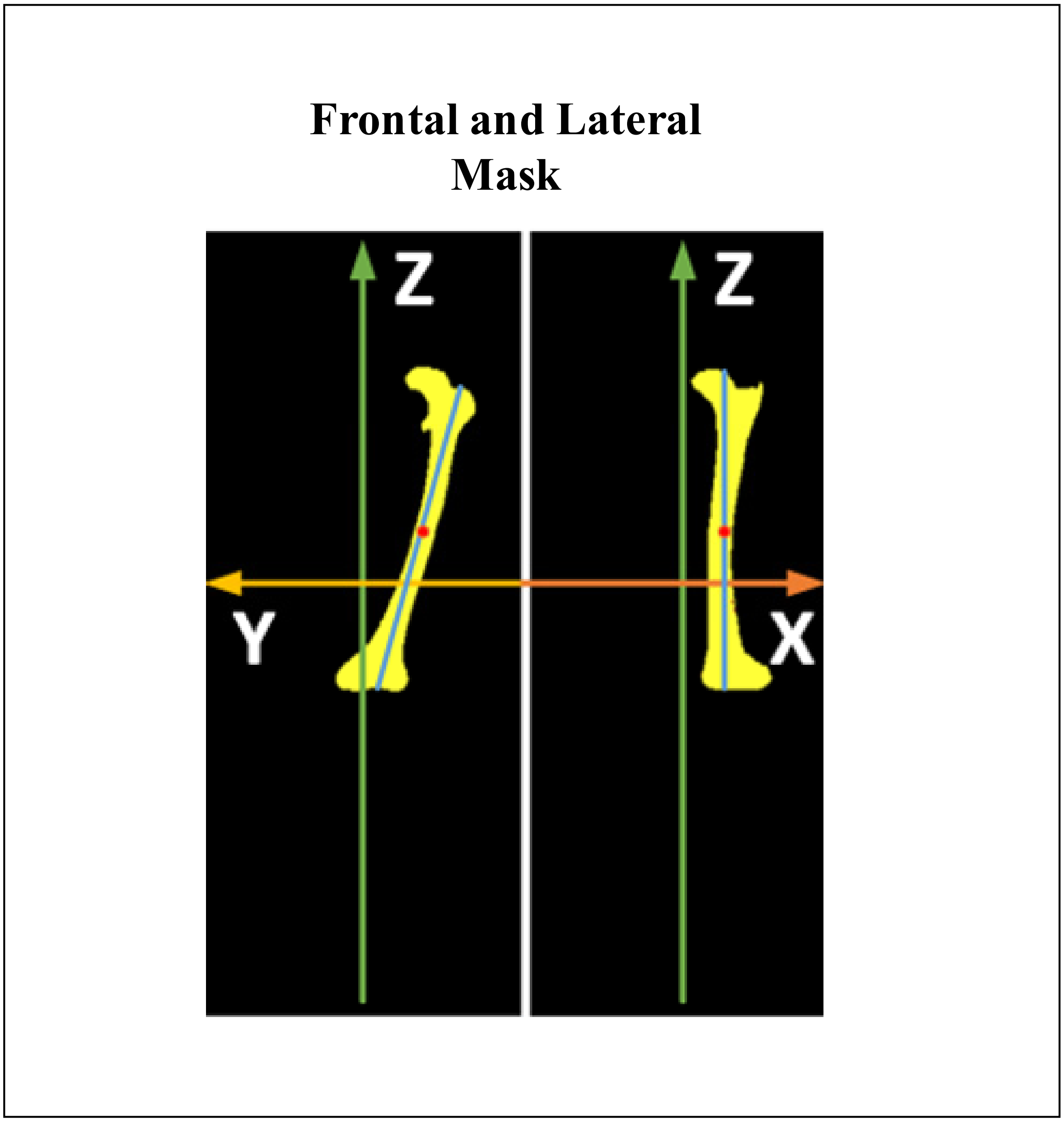


Figure 4.8 Illustration of applying PCA on frontal and lateral mask of the femur

4.3.3.1 Computing 3D translations (t_x, t_y, t_z)

Given (C_F^{2D}, C_L^{2D}), the corresponding 3D center position of the target 3D bone structure (F), C_{Bone}^{3D}, is found as the midpoint of the shortest line segment joining the frontal and lateral X-ray projection paths (P_F^{3D}, P_L^{3D}) (Chaibi et al., 2012 ; Yu et al., 2016). Firstly, the center of the generic 3D model, C_g^{3D}, is calculated as Eq. (4.6):

$$C_g^{3D} = \left(\frac{\sum_{i=1}^{v} x_i}{v}, \frac{\sum_{i=1}^{v} y_i}{v}, \frac{\sum_{i=1}^{v} z_i}{v}\right) \tag{4.6}$$

Then, the 3D translations are computed as 3D displacements between C_{Bone}^{3D} and C_g^{3D}, as Eq. (4.7):

$$(t_x, t_y, t_z) = C_{Bone}^{3D} - C_g^{3D} \tag{4.7}$$

where, v is the number of vertices in (M_g), and (x_i, y_i, z_i) is the 3D coordinates of the i^{th} vertex.

4.3.3.2 Computing rotations (r_θ) and (r_α)

Given $(Mask_F, Mask_L)$, we compute the eigenvectors and eigenvalues of the covariance matrix of the 2D coordinate positions of the mask pixels, as Eqs. 4.8 and 4.9, (Salvi et al., 2007). The eigenvectors (EV) with the highest eigenvalues show the principal directions with the largest variances on $(Mask_F, Mask_L)$, (Figure 4.9). Eqs. 4.10 and 4.11 allow to compute the principal directions' angles (R_F^{2D}, R_L^{2D}). Given the current orientation of $\left(M_g\right)$ at the origin of the 3D referential system (0,0,0), as R_g^θ, R_g^α, and $\left(R_g^\beta = 0\right)$, the two rotation angles (r_θ) and (r_α) to align $\left(M_g\right)$ into (F), are computed as Eq. (4.10) and Eq. (4.11), respectively. (Figure 4.9) shows computed two rotation angles (r_θ) and (r_α).

$$R_F^{2D} = tan^{-1}\left(\left|\frac{\left(EV(COV(Mask_F))_{(Z)}\right)}{\left(EV(COV(Mask_F))_{(Y)}\right)}\right|\right) \tag{4.8}$$

$$R_L^{2D} = tan^{-1}\left(\left|\frac{\left(EV(COV(Mask_L))_{(Z)}\right)}{\left(EV(COV(Mask_L))_{(X)}\right)}\right|\right) \tag{4.9}$$

$$(r_\theta) = R_F^{2D} - R_g^{\theta} \tag{4.10}$$

$$(r_\alpha) = R_L^{2D} - R_g^{\alpha} \tag{4.11}$$

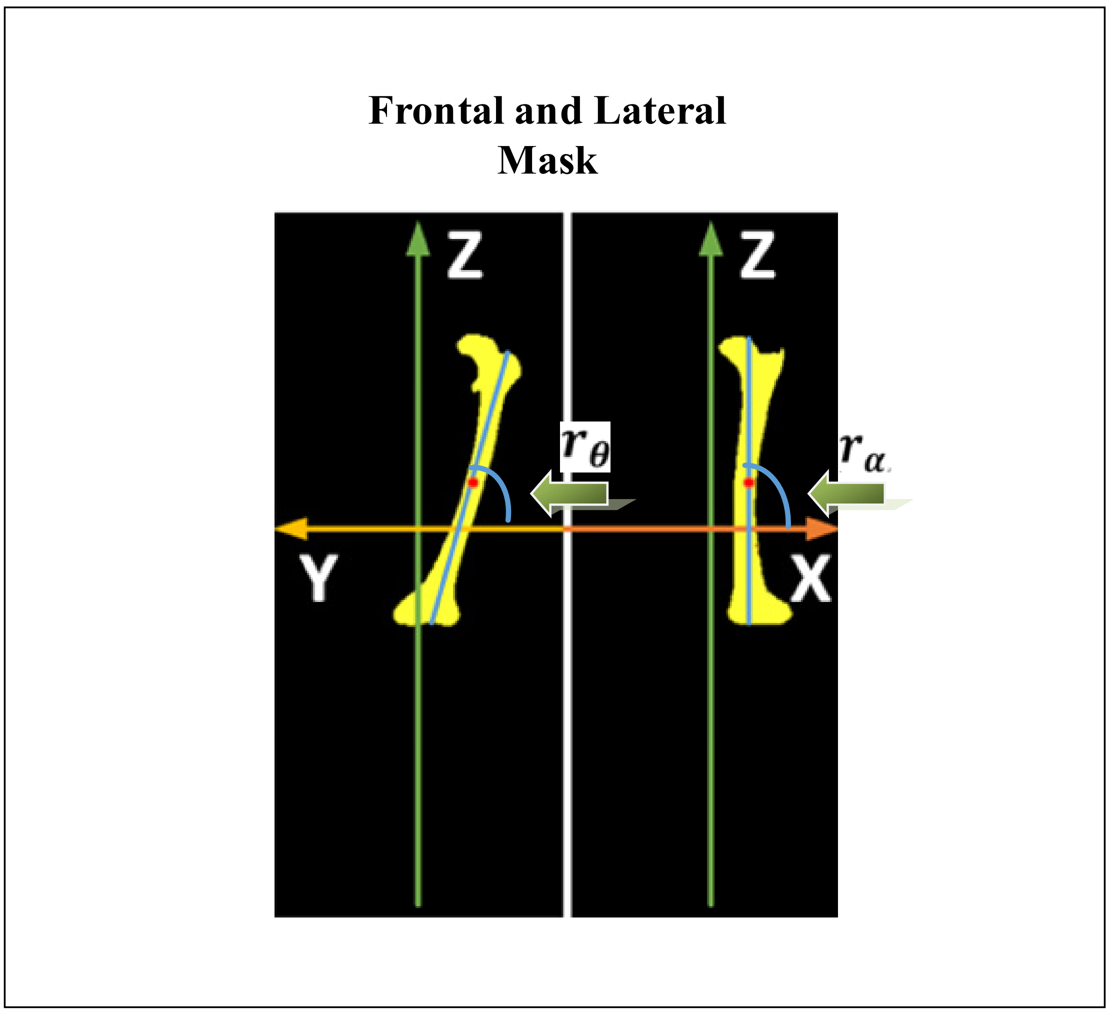


Figure 4.9 Illustration of two computed rotation angles using PCA on bi-planar masks

4.3.3.3 Computing 3D isotropic scaling ratio (s)

We align (M_g) into (F) by computed two rotations (r_θ) and (r_α), and 3D translations (t_x, t_y, t_z) to obtain the transformed model as (M_g^{rt}), Eq. (4.12):

$$(M_g^{rt}) = (r_\alpha) \cdot \left((r_\theta) \cdot (M_g) \right) + (t_x, t_y, t_z) \tag{4.12}$$

Then, the isotropic scaling (s) between (M_g^{rt}) and (F) is computed as Eq. (4.13):

$$(s) = \frac{L_F}{L_{M_g^{rt}}} \tag{4.13}$$

Where, L_F and L_{M_g} are the 3D Euclidean distances between the lowest and the highest 3D point of (F) and the aligned model (M_g^{rt}), respectively, along the vertical axis of the femur bone and the corresponding mask (Figure 4.10). Then, we apply the computed isotropic scaling (s) on (M_g^{rt}) to obtain the transformed model (M_g^{srt}), as Eq. (4.14):

$$(M_g^{srt}) = (s) \cdot (M_g^{rt}) \tag{4.14}$$

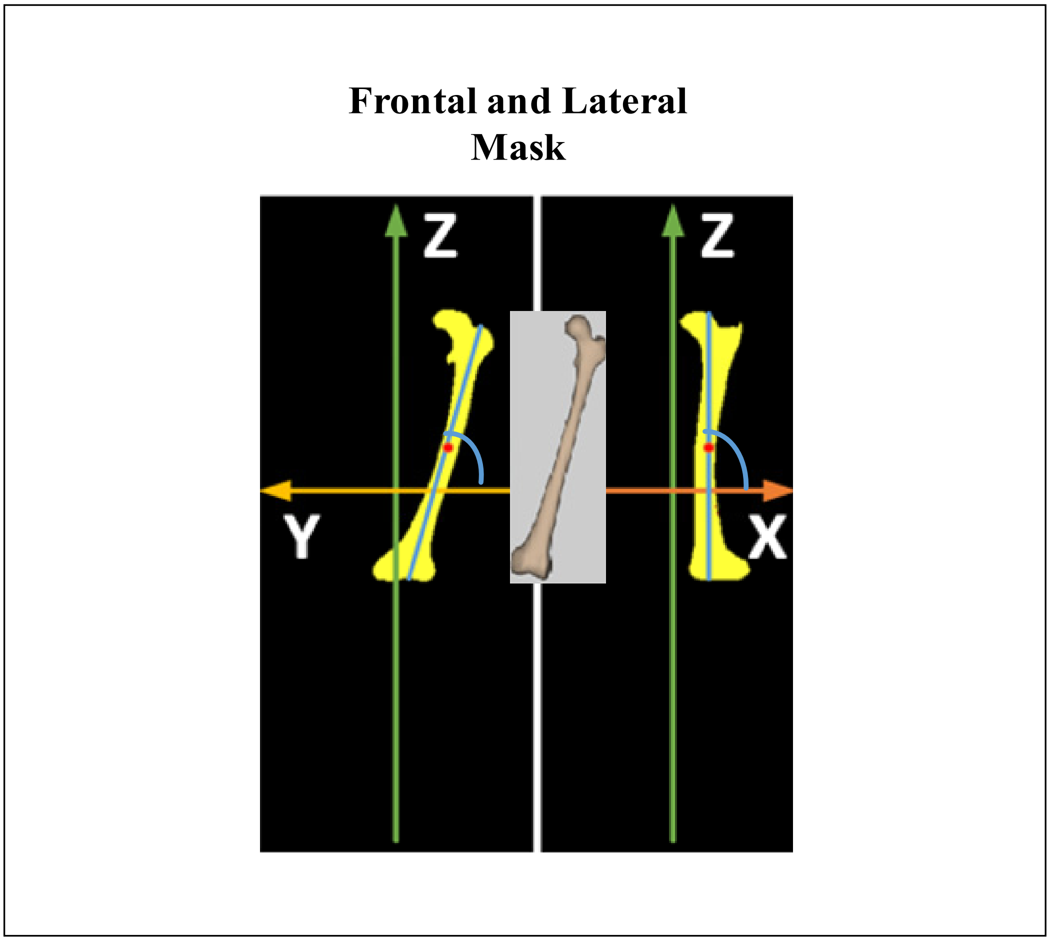


Figure 4.10 Computing the 3D Euclidean distances between aligned 3D model and masks along the vertical axis of the femur

4.3.4 CNN-based registration

In 3D/2D similarity registration, we train four twelve-channel CNN-based regression models: $CNN^1_{(r_\beta)}$, $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$, with the same architecture as described in 4.2.2. $CNN^1_{(r_\beta)}$ is trained in the coarse registration step to estimate (r_β), and then in the fine registration step $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$ are trained to estimate transformation parameters' residuals, $\delta T(\delta t_x \delta t_y, \delta t_z, \delta r_\theta, \delta r_\alpha, \delta r_\beta)$.

4.3.5 CNN-based registration to estimate (r_β)

In the coarse registration (Figure 4.7), after estimation of the six parameters $(t_x t_y, t_z, r_\theta, r_\alpha, s)$ via PCA-based registration, we train a twelve-channel CNN-based regression model, $CNN^1_{(r_\beta)}$, as described in section 4.2.2 (Figure 4.4), to estimate out-of-plane rotation (r_β). The input of the twelve-channel CNN-based regressor is 12 LIRs, where each channel of the CNN corresponds to each LIR. To prepare the input images for $CNN^1_{(r_\beta)}$, the Ray Casting method (Roth D. S., 1982) is used to render 2D bi-planar digitally reconstructed radiographs (DRRs) of the transformed generic 3D model (M_g^{srt}). Then, local intensity differences (LIRs) between (DRRs) of the transformed generic 3D model, (M_g^{srt}) as source, and the patient's 2D bi-planar radiographs, as reference images, are calculated. The input LIRs generation process will be explained, in details, by section 4.3.6. After regression, (M_g^{srt}) is rotated by the estimated (r_β), to obtain (M_g^T). Afterwards, the transformed generic 3D model (M_g^T) is used as input of the fine registration.

4.3.6 Generation pose-invariant local intensity residuals (LIRs)

In the coarse and the fine registration, the input of the multi-channel CNN-based regression models are local intensity residuals (LIRs). The number of channels corresponds to the number of input LIRS. The number of input LIRs varied in the coarse and the fine registration step. In the coarse registration step, the input of the twelve-channel CNN-based regressor, $CNN^1_{(r_\beta)}$, is 12 LIRs, where each channel of the CNN corresponds to each LIR. In the fine registration step, the input of multi-channel CNN-based regression models: $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$, is six frontal, six lateral, and 12 bi-planar LIRs, respectively. Each channel of the multi-channel CNN corresponds to each input LIR. In the coarse registration, the input local intensity residuals (LIRs) are local intensity differences between 2D bi-planar digitally reconstructed radiographs (DRRs) of (M_g^T), after the PCA-based registration, as source images, and the real 2D bi- planar radiographs of the patient, as reference images. After the coarse registration step, in the fine registration, the input LIRs are calculated between 2D DRRs of the coarsely registered 3D model and the patient 2D radiographs. The LIRs are computed by subtracting the pixels' intensity value, inside the extracted local patches, of 2D bi-planar DDRs from 2D bi-planar radiographs. The extracted local patches are centered at 2D projection of pose-invariant 3D points on 2D bi-planar DDRs and 2D bi-planar radiographs, which lead to create pose-index features (Miao et al., 2016). Figure 4.11 shows six frontal and six lateral pose-invariant LIRS, after coarse registration. In Figure 4.11, the first row shows 2D projections of the six frontal and the six lateral pose-invariant 3D points on the left femur. The second and third rows show the extracted six frontal and six lateral LIRS, respectively.

Six frontal and six lateral pose-invariant local patches are centered at 2D projections of the 12 selected pose-invariant 3D points on frontal and lateral silhouettes of (M_g). Figure 4.11 shows six frontal and six lateral pose-invariant LIRS. The selection of the pose-invariant 3D points is very important since we need to create pose-invariant intensity features in order to make the registration results insensitive to initial transformation parameters (t) and highly sensitive to transformation parameter residual (δt) by estimating the same 3D rigid transformation parameters from different initial position (Miao et al., 2016). The input LIRs of each multi-channel CNN-based regression models: $CNN^1_{(r_\beta)}$, $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$, are computed in the same size of (300×300) pixels. The size of the all LIRs is the same as (300×300) pixels and is selected empirically which is large enough to surround the important intensity features and shape features such as the curves and lines around the corresponding pose-invariant center point.

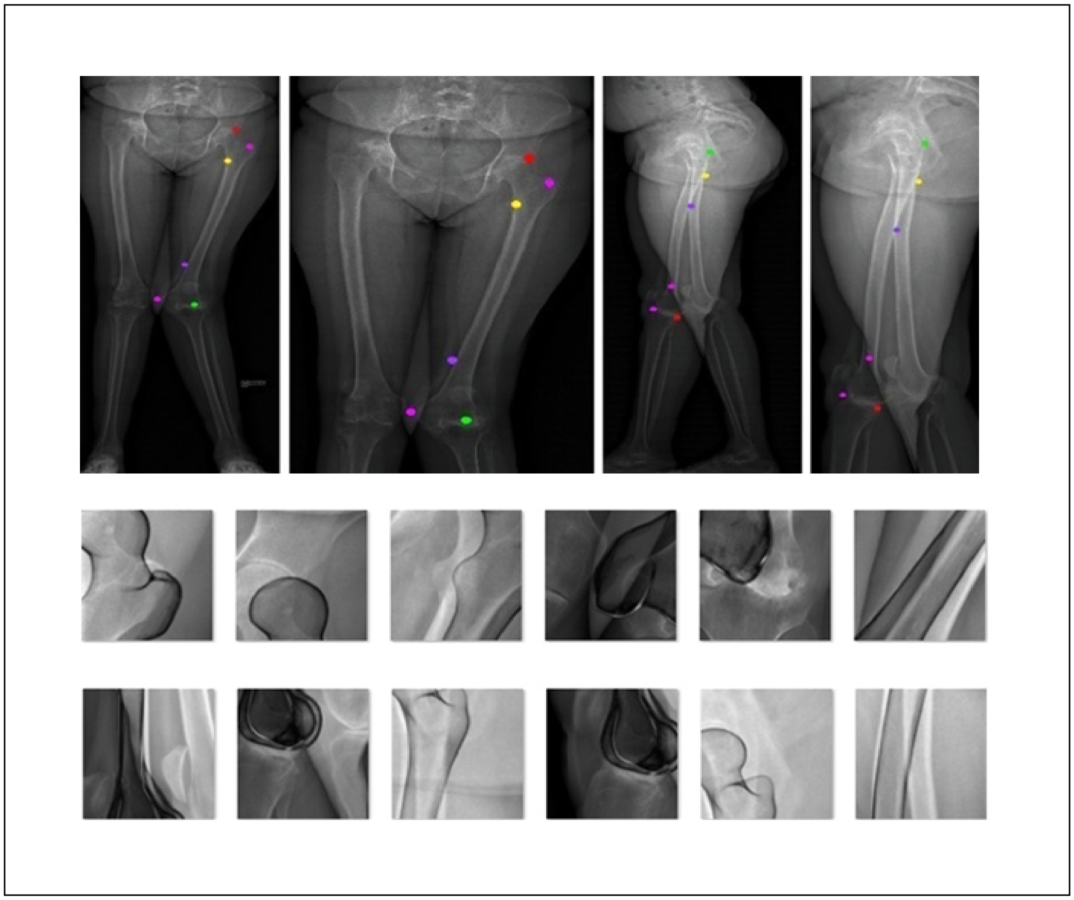


Figure 4.11 2D projections of six pose-invariant 3D points on femur and LIRs

To select the pose-invariant 3D points and corresponding local patches, we render- empirically determined- 100 DRRs of (M_g^T), in 10 random transformation parameters (t), as sources images, and 10 random displacements $(t + \delta t)$ for each (t), as reference images. Hence, for 10 (t) and 10 (δt) for each (t), we end up with 100 different $(t + \delta t)$ and 100 DRRs. The pose-invariant 3D points are more sensitive to (δt) and selected based on a variance-based measurement (Miao et al., 2016). For each 3D point i on the silhouettes of (M_g), we compute LIRs between (t) and $(t + \delta t)$. The sensitivity measurement is Eq. (4.15):

$$Sensitivity_i = \frac{\{S_{\delta t}\}_i}{\{S_t\}_i} \tag{4.15}$$

where, $\{S_t\}_i$ and $\{S_{\delta t}\}_i$ measure the sensitivity of LIRs to (t) and (δt), respectively. We select the six most sensitive pose-invariant 3D points, with the highest $Sensitivity_i$, on frontal and lateral silhouette, independently (Youn et al., 2017). Algorithm 4.1 presents the process of sensitivity measurement to select the six most sensitive pose-invariant 3D points.

Algorithm 4.1 Sensitivity measurement

Input: An array of 3D points, $\{{P_i}^{3D}\}$, on the silhouette of (M_g)

Output: Six most sensitive 3D points of (M_g)

For $i = 1\ to\ n\ , i = 1, \dots, n\ \in \{{P_i}^{3D}\}$

    For a=1 to 10

        For b=1 to 10

            For c=1 to number of pixels inside the local patch $\{L_t\}_{iab}$

                Compute $\{\{IL_t\}_{ciab} - \{IL_{t+\delta t}\}_{ciab}\}_i$

            End

        End

        Compute $mean_b(\{IL_{t+\delta t}\}_{ciab}\}_i)$

    End

    Compute $mean_a(\{\{IL_t\}_{ciab}\}_i)$

    Compute $\{S_t\}_i$ = $mean((\{\{IL_t\}_{ciab} - \{IL_{t+\delta t}\}_{ciab}\}_i$-$(mean_a(\{\{IL_t\}_{ciab}\}_i))$^2)

    Compute $\{S_{\delta t}\}_i$ = $mean((\{\{IL_t\}_{ciab} - \{IL_{t+\delta t}\}_{ciab}\}_i - (mean_b(\{IL_{t+\delta t}\}_{ciab}\}_i))$^2)

    Compute $Sensitivity_i$

    $Array\_Sensitivity$ ← Add-to-Array ($Sensitivity_i$)

End

$Array\_Sensitivity$ ← Sort-Descended ($Array\_Sensitivity$)

$Array\_Sensitivity$ ← Remove overlapping more than 20% ($Array\_Sensitivity$)

$Most\ Sensitive\ 3D\ Points$ ← Select six points of highest sensitive ($Array\_Sensitivity$)

Return ($Most\ Sensitive\ 3D\ Points$)

The number of the six most sensitive 3D points is enough and efficient since the location and the local intensity around them shows the more important features which are useful to find 3D transformation parameters. The six most sensitive 3D points are located in the femoral regions, where contribute much in providing the best fit in the femoral shape deformation process (Chaibi et al., 2012 ; Youn et al., 2017). The six most sensitive pose-invariant 3D points are located on femoral head, greater trochanter, lesser trochanter, lateral and medial condyles (Figure 4.12).

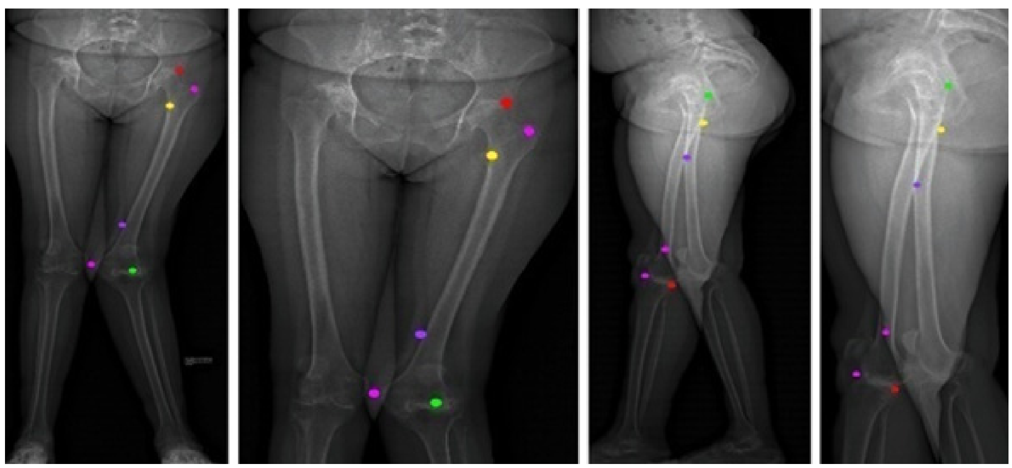


Figure 4.12 2D projections of six frontal (left) and lateral (right) pose-invariant 3D points on the left femur

4.3.7 CNN-based fine registration

After the coarse registration, in the fine registration step (Figure 4.13) to refine the 3D pose parameters, transformation parameters' residuals, $\delta T(\delta t_x \delta t_y, \delta t_z, \delta r_\theta, \delta r_\alpha, \delta r_\beta)$, are estimated through three multi-channel CNN-based regression models: $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$, with the same structure as described in section 4.2.2 (Figure 4.4). In the fine registration, we divide $\delta T(\delta t_x \delta t_y, \delta t_z, \delta r_\theta, \delta r_\alpha, \delta r_\beta)$ into three groups in order to define three different CNN models. The input of multi-channel CNN-based regression models: $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$, is six frontal, six lateral, and 12 bi-planar LIRs, respectively. Hence, $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$ are six-channel, six-channel, and twelve-channel CNN-based regressors, where the number of the channels corresponds to the number of input LIRs. The output of $CNN^2_{(\delta t_y, \delta r_\theta)}$, $CNN^3_{(\delta t_x, \delta r_\alpha)}$, and $CNN^4_{(\delta t_z, \delta r_\beta)}$ are $(\delta t_y, \delta r_\theta), (\delta t_x, \delta r_\alpha)$, $(\delta t_z, \delta r_\beta)$ from six frontal, six lateral, and 12 bi-planar input LIRs, respectively. The input of the fine registration step includes: (M_g^T), 2D bi-planar DRRs of (M_g^T), and 2D bi-planar radiographs of the femur. We refine six transformation parameters, and then transform (M_g^T) with the refined 3D pose parameters. In the fine registration, the isotropic scaling ratio (s) is not refined since in the coarse registration step we achieve sub-percentage precision.

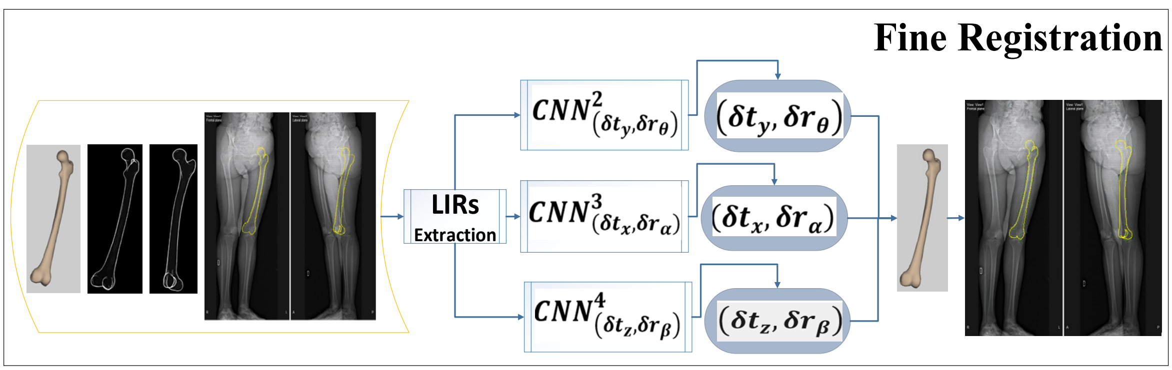


Figure 4.13 The workflow for fine registration step

4.4 Automatic 3D/2D non-rigid registration

4.4.1 Introduction

Automatic 3D/2D non-rigid registration (Figure 4.2, bottom row) includes two steps:

1) Local 3D shape deformation (U_{Shape});
2) Local 3D scale deformation (U_{Scale}).

To deform the local shape of the registered 3D femur model, CNN-based regression models are trained to find 3D displacements of a small number of 3D handles pre-defined on the femur. Following the computation of the new positions of 3D handles, we apply MLS deformation to obtain a 3D model better adjusted to the radiographs without any user interactions. To personalize the local 3D scale corresponding to each handle, CNN-based regression models are trained to estimate the local 3D scale ratios corresponding to the said handles. Then, we use the MLS method to compute 3D rescaling fields and extend the scales to the entire 3D femur.

4.4.2 Automatic 3D shape deformation ($\boldsymbol{U_{Shape}}$)

To perform a personalized and as-rigid-as-possible local 3Dbone deformation without undesirable distortion, we propose a fast and automatic handle-based 3D/2D non-rigid registration algorithm combined with MLS deformation (Chaibi et al., 2012 ; Cresson et al., 2010). In U_{Shape}, first, we train CNN-based regression models to estimate the 3D displacements of a small set of 17 handles of the generic 3D model (4.2.1, Figure 4.5). Given the new positions of the handles, we apply MLS deformation (4.2.1.1) without any user interventions.

4.4.2.1 CNN-based handles' 3D displacement estimation

To deform the 3D model $\left(M_g^T\right)$, we estimate 3D displacements of the set of 17 handles, $\boldsymbol{CP}$ (Figure 4.5). A set of 17 CNN-based regressors, $\boldsymbol{CNNs^{Shape}} = \left\{CNN_i^{Shape} \middle| i = 1, \cdots, 17\right\}$, are independently trained to directly estimate the handles' 3D displacements between the current and target position, $\boldsymbol{\Delta}_{\text{xyz}} = \left\{(\delta x, \delta y, \delta z)_i \middle| i = 1, \dots, 17\right\}$. The input of each CNN_i^{Shape} is local intensity (LIs) (section 4.3.3) inside the bi-planar local patches extracted from patient's 2D bi-planar radiographs (I_F, I_L) at a size of 200×200 pixels. The input LIs of CNN_i^{Shape} are centered at 2D frontal and lateral projections of the 3D handles, ${CP_i}^T$, (Figure 4.14 and 4.15) transformed using estimated similarity transformation T, in section 4.2. Each

CNN_i^{Shape} (Figure 4.5) learns a map function f_1, as Eq. (4.16), between the input LIs and corresponding 3D displacements $(\delta x, \delta y, \delta z)_i$:

$$f_1\left(LIs\left(CP_i^T, (I_F, I_L)\right), (M_g^T)\right)_i = (\delta x, \delta y, \delta z)_i \tag{4.16}$$

Afterwards, $\boldsymbol{CP}^{'}$ is computed as a set of new positions for transformed 17 handles by T,$\boldsymbol{CP}^{T}$, as Eq. (4.17):

$$\boldsymbol{CP}^{'} = \boldsymbol{CP}^{T} + \boldsymbol{\Delta}_{\mathrm{xyz}} \tag{4.17}$$

Each CNN_i^{Shape} comprises two channels; the top channel deals with frontal local patch, while the bottom one covers lateral local patch (Figure 4.5). This allows to directly estimating 3D displacements. The input and the structure of CNN_i^{Shape} are designed to obtain 3D displacements as output. Since it is important to compute 3D displacements of the MLS handles at the origin of the 3D referential of the EOS® cabin system $(0,0,0)$, the inverse of the 3D/2D similarity transformation, T^{-1}, is applied in the definition of the Euclidean loss function, as Eq. (4.18):

$$\psi = \frac{1}{n}\sum_{j=1}^{n}\left\| T_j^{-1}\left((\delta x, \delta y, \delta z)_j^{Target}\right) - T_j^{-1}\left((\delta x, \delta y, \delta z)_j\right)\right\|_2^2 \tag{4.18}$$

Where n is the number of samples and $T_j^{-1}\left((\delta x, \delta y, \delta z)_j^{Target}\right)$ and $T_j^{-1}\left((\delta x, \delta y, \delta z)\right)$ are the known target and estimated 3D displacements for the j^{th} training sample, respectively.

4.4.2.2 Automatic 3D shape deformation

The automatic CNN-based computation of $\boldsymbol{CP}^{'}$ allows to apply the MLS 3D shape deformation without any user annotations and interactions. Following the computation of new positions for 17 3D handles, $\boldsymbol{CP}^{'}$, Eq. (4.18), to obtain an as-rigid-as-possible 3D shape deformation, we use the MLS method, described in section 4.2.1.1, to estimate regularized 3D displacements over all vertices of $\left(M_g^T\right)$. The deformed model is called $\left(M_g^{\widehat{D\circ}(T)}\right)$.

4.4.3 Automatic 3D scale deformation ($\boldsymbol{U_{Scale}}$)

Following ($CNNs^{Shape} + MLS$), to personalize the 3D scales of $\left(M_g^{\widehat{D}\circ(T)}\right)$ according to the target bone (F), we develop an automatic 3D scale deformation algorithm. $\mathrm{U_{Scale}}$ combines CNN-based regression models with a regularized MLS-based local 3D scale deformation to drive a fully automatic 3D scale deformation. The CNN-based regressors estimate local scale ratios corresponding to the 17 displaced 3D handles, which are required to perform an as-rigid-as-possible MLS-based local 3D scale deformation without any user interventions.

4.4.3.1 CNN-based handles' 3D scale ratios estimation

To adjust the scale of the handles on $\left(M_g^{\widehat{D}\circ(T)}\right)$, we develop 17 CNN-based regressors, $\boldsymbol{CNNs^{Scale}} = \left\{CNN_i^{Scale} \middle| i = 1, \cdots, 17\right\}$, with the same architecture as $\boldsymbol{CNNs^{Shape}}$, but with a different cost function, as Eq. (4.19):

$$\psi = \frac{1}{n}\sum_{j=1}^{n} \left\| \left(\delta s_x, \delta s_y, \delta s_z\right)_j^{Target} - \left(\delta s_x, \delta s_y, \delta s_z\right)_j \right\|_2^2 \tag{4.19}$$

where n is the number of samples and $\left(\delta s_x, \delta s_y, \delta s_z\right)_j^{Target}$ and $\left(\delta s_x, \delta s_y, \delta s_z\right)_j$ are the known target and estimated scale ratio for the j^{th} training sample, respectively. The $\boldsymbol{CNNs^{Scale}}$ are independently trained to estimate the 3D scale ratios $\mathbf{SR}_{\mathrm{xyz}} = \left\{\left(\delta s_x, \delta s_y, \delta s_z\right)_i \in \mathbb{R} \middle| i = 1, \dots, 17\right\}$ of the 17 displaced 3D handles, $\boldsymbol{CP}^{'}$. Among the set of 17 3D handles, $\boldsymbol{CP}^{'}$, $CP_1^{'}$ to $CP_{12}^{'}$ are point handles with uniform scale $\left(s_x = s_y = s_z\right)_i$, and the corresponding regressors $\{CNN_i^{Scale}|i=1,\cdots,12\}$ have one output in the last layer. CP'_{13} to CP'_{17} ($\{CNN_i^{Scale}|i=13,\cdots,17\}$) are the center points of the spline handles with non-uniform scale, which therefore have three outputs. Each CNN_i^{Scale} learns a direct map function f_2 between the input 2D bi-planar (LIRs), section 4.3.3, and the output scale ratios, as Eq. (4.20):

$$f_2\left(LIRs\left(CP'_i,(DRR_F,DRR_L),(I_F,I_L)\right),\left(M_g^{\widehat{D\circ}(T)}\right)\right)_i=\left(\delta s_x,\delta s_y,\delta s_z\right)_i \quad (4.20)$$

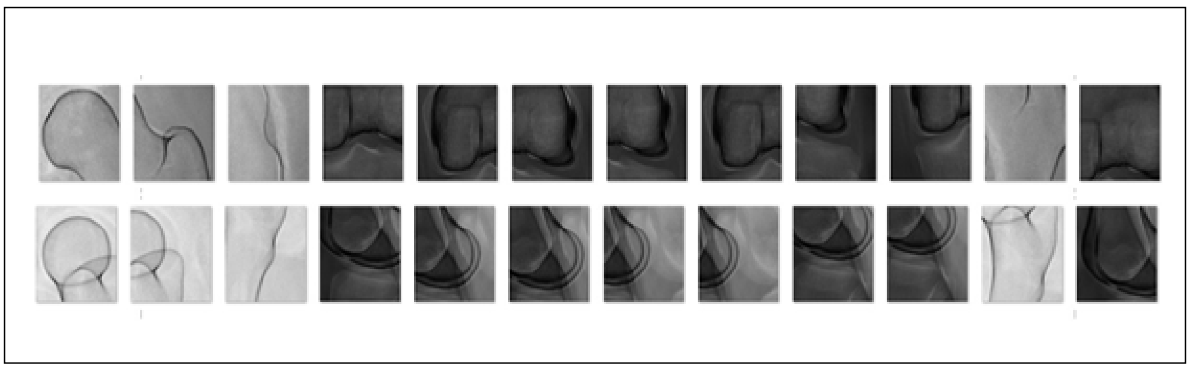


Figure 4.14 Illustration of 12 frontal LIRs (top) and lateral LIRs (bottom) centered at 2D projection of 3D handles with uniform scaling CP'_1 to CP'_{12}

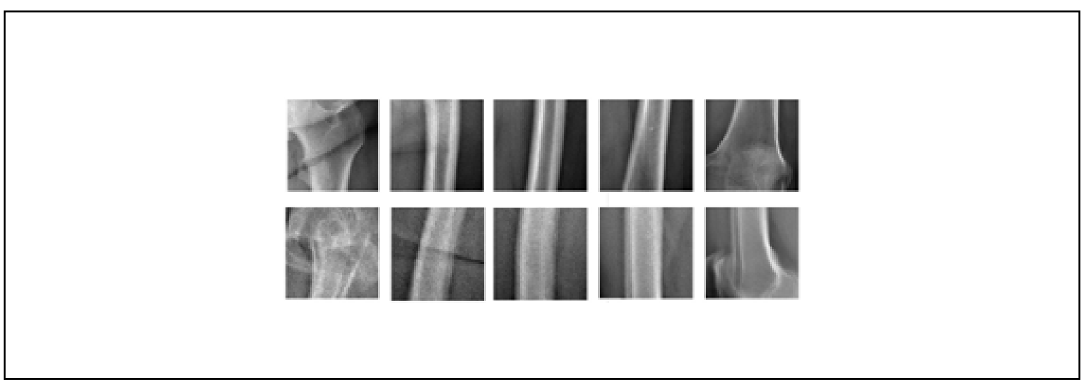


Figure 4.15 Illustration of 12 frontal LIRs (top row) and lateral LIRs (bottom row) centered at 2D projections of 5 3D handles with non-uniform scaling CP'_{13} to CP'_{17}

Where the input LIRs are local intensity differences, inside extracted local patches, between bi-planar digitally reconstructed radiographs (DRRs) of $\left(M_g^{\widehat{D}\circ(T)}\right)$, as source (DRR_F, DRR_L), and patients' bi-planar radiographs as reference (I_F, I_L). The extracted local patches (Figure 4.14 and 4.15) are centered at 2D projections of the 17 displaced 3D handles, $\boldsymbol{CP}'$, on both frontal and lateral projections, at a size of 200×200 pixels. The size of the local patch is determined empirically by considering the fact that the local patch is large enough to contain the curves and shape of the region corresponding to each 3D handle.

4.4.3.2 Automatic 3D scaling

Following a CNN-based 3D scale ratios estimation, we automatically apply scaling over all vertices to personalize the scales of $\left(M_g^{\widehat{D}\circ(T)}\right)$. The sum of weighted CNN-based local 3D scale ratios of the 17 handles over the sum of the computed weights, w_i,is computed to obtain the average of scale ratios over all handles, $\overline{SC}_{Local}$, Eq. (4.21). Then, the average of SC_{Global}, described in Section 4.3.2.1 (Eqs. 4.2 and 4.3), and $\overline{SC}_{Local}$, are calculated to obtain an optimal local scale $\widehat{SC}_{V_i}$, Eq. (4.22), as (Chaibi et al., 2012):

$$\overline{SC}_{Local} = \frac{\sum_i w_i\left(\delta s_x, \delta s_y, \delta s_z\right)_i}{\sum_i w_i}.\ i = 1, \cdots, 17 \tag{4.21}$$

$$\widehat{SC}_{V_i} = \frac{(sc_{Global} + \overline{SC}_{Local})}{2} \tag{4.22}$$

Ultimately, each vertex $V_i \in \left(V_g^{\widehat{D}\circ(T)}\right)$ is transformed with the computed optimal scale $\widehat{SC}_{V_i}$ to extrapolate 3D scale deformation over $\left(M_g^{\widehat{D}\circ(T)}\right)$ and obtain a personalized (deformed and rescaled) 3D bone model as $\left(M_g^{\hat{S}\circ\left(\widehat{D}\circ(T)\right)}\right)$.

CHAPITRE 5

EXPERIMENTS AND RESULTS

5.1 Experimental data and setup

5.1.1 Experimental data and setup of 3D/2D similarity registration

5.1.1.1 2D Bi-planar Radiographs

After ethical approvals by the ethics committees of the University of Montreal Hospital Center (CHUM) and École de technologie supérieure (ÉTS, Montréal, Canada), a set of 85 patients' EOS® 2D bi-planar radiographs is retrospectively recovered with two different fields of view, full body and whole lower limbs, in two patients' orientations, $0^{\circ}/90^{\circ}$ and $45^{\circ}/45^{\circ}$, which are acquired by the low dose EOS® imaging system. An expert, via a semi-automatic commercial software tool, SterEOS (Chaibi et al., 2012) as an state-of-the-art (SOTA) method, generates patient-specific fuzzy gold standard (Jannin et al., 2002) 3D reconstructed models of the left femur (M_{FGS}) corresponding to the 85 patients. For each case, the fuzzy gold standard six 3D pose parameters and isotropic scaling are derived by solving a least-square regression between $\left(M_g\right)$, as source model, and the corresponding patient-specific(M_{FGS}), as reference, using a singular value decomposition (SVD) method (Arun, et al., 1987). From this set, 30 are used for training, 30 for validation, and 25 for unseen testing.

5.1.1.2 Digitally reconstructed radiographs (DRRs)

The 2D bi-planar DRRs are rendered from the transformed 3D model of the left femur $\left(M_g^T\right)$, presented in section 4.3, using the Ray Casting method (Roth, 1982) via a home-made algorithm developed in C++, and running on an Intel® Core CPU. This Ray Casting method simulates the same geometry of the EOS® cabin system for 2D bi-planar projection. We render 85 bi-planar DRRs of (M_g^T), corresponding to 85 patients (training, validation, and test set).

5.1.1.3 Training data

In 3D/2D similarity registration (U_{Pose}), four CNN-based regression models ($CNN^1_{(r_\beta)}$, $CNN^2_{(\delta t_y,\delta r_\theta)}$, $CNN^3_{(\delta t_x,\delta r_\alpha)}$, and $CNN^4_{(\delta t_z,\delta r_\beta)}$) are trained to estimate (r_β), $(\delta t_y, \delta r_\theta)$, $(\delta t_x, \delta r_\alpha)$, and $(\delta t_z, \delta r_\beta)$, respectively. The training data set comprises 30 bi-planar radiographs acquired from 15 patients with $0^\circ/90^\circ$ and 15 patients with $45^\circ/45^\circ$ orientations in the EOS® cabin system with full body or whole lower limb fields of view, as reference images. In the coarse registration, for each reference bi-planar radiograph of 30 patients, we generate synthetic 2D bi-planar DRRs of (M_g^T) in the corresponding fuzzy gold standard 3D pose parameters with $(r_\beta = 0)$, as source images. In the fine registration, for each of 30 patients, we generate 10 random perturbations around the corresponding fuzzy gold standard 3D pose parameters. The perturbations follow a uniform distribution inside the range of ± 15 mm for (t_x, t_y, t_z), $\pm 15°$ for (r_θ) and (r_α), and $\pm 45°$ for (r_β). For 30 patients in the training set, we totally generate 300 2D bi-planar DRRs of the transformed 3D model by the generated random perturbations, (M_g^T), in by the generated random perturbations. For each patient, we extract six frontal and six lateral local patches leading to 360 2D local pose-invariant patches, as described in section 4.3.6. In total, for all 30 patients, we end up with 3600 2D local pose-invariant patches. The four convolutional neural network models are implemented in a homemade Tensor-flow framework on a GeForce TITANX GPU. The DRRs are rendered by the Ray Casting method (Roth, 1982) via an algorithm developed in C++ and running on an Intel CPU, in the laboratory LIO.

5.1.1.4 Validation data set and scheme

The validation scheme carries out a thorough investigation in three different experiments to compare the performance and accuracy of simulated and real applications. The first experiment conducts a 3D/2D similarity registration from (M_{FGS}), as a simulated model, into the corresponding fuzzy gold standard masks ($Mask_F^{FGS}, Mask_L^{FGS}$), as simulated masks, to validate feasibility of proposed methodology without possible errors introduced by the generic 3D model and the CNN-based segmented masks. The corresponding fuzzy gold standard masks are obtained by projecting the 3D silhouette of the fuzzy gold standard 3D models of the femur on 2D bi-planar radiographs of patients. The second experiment registers (M_g) into ($Mask_F^{FGS}, Mask_L^{FGS}$), to validate the effects of the generic 3D model as a source. The second experiments reveal the impact of the personalized 3D femur's shape on the estimated 3D pose parameters accuracy. Ultimately, the third experiment investigates a real application; 3D/2D similarity registration of (M_g) into CNN-based ($Mask_F, Mask_L$)to validate the efficacy, robustness, and accuracy of the proposed method. The fuzzy gold standard masks ($Mask_F^{FGS}, Mask_L^{FGS}$) are more accurate than the CNN-based segmented masks and the difference between the second and third experiment shows the effects of the segmentation error on the 3D/2D similarity registration accuracy.

5.1.1.5 Evaluation metric

In both coarse and fine registrations, Mean and Standard Deviation of Absolute Pose Errors $(\mathrm{MAE} \pm \mathrm{STD})$ between estimated and fuzzy gold standard 3D pose parameters and isotropic scaling ratio are computed over evaluation cases.

5.1.2 Experimental setup and data of 3D/2D non-rigid registration

5.1.2.1 2D bi-planar radiographs

After ethical approvals by the ethics committees of the University of Montreal Hospital Center (CHUM) and École de technologie supérieure (ÉTS, Montréal, Canada), a set of 2D bi-planar radiographs of 85 patients are retrospectively recovered, in which the whole femur is presented in bi-planar radiographs. The set of 2D bi-planar radiograph are acquired from various fields of view, including the whole lower limbs and the full body, with two different patient orientations, $0^{\circ}/90^{\circ}$ and $45^{\circ}/45^{\circ}$, by the low dose and slot-scanning EOS® system (EOS® Imaging, Paris, France). We use a total of 70 patients (56 for training and 14 for validation). 15 unseen patients' 2D bi-planar radiographs are then used for testing. From this set, nine and six bi-planar radiographs are acquired with $0^{\circ}/90^{\circ}$ and $45^{\circ}/45^{\circ}$ patients' orientations in the EOS® cabin system, respectively. From nine 2D bi-planar radiographs with $0^{\circ}/90^{\circ}$ orientation, two and seven are in full body and whole lower limbs fields of view, respectively. All of the six 2D bi-planar radiographs with $45^{\circ}/45^{\circ}$ orientation are in whole lower limbs fields of view.

5.1.2.2 Digitally reconstructed radiographs (DRRs)

The 2D bi-planar DRRs are rendered from the transformed and then deformed 3D model of the left femur $\left(M_g^{\widehat{D}\circ(T)}\right)$, presented in section 4.4, using the Ray Casting method (Roth, 1982) described in section 5.1.1.2. We render 85 bi-planar DRRs of $\left(M_g^{\widehat{D}\circ(T)}\right)$, corresponding to 85 patients (training, validation, and test set).

5.1.2.3 Fuzzy gold standard personalized 3D models

To evaluate the performance of the 3D/2D registration, we compare the accuracy of the personalized 3D shape of the femur, 3D position, and local scale of 17 handles with corresponding personalized fuzzy gold standard (Chaibi et al., 2012) 3D shape, fuzzy gold standard 3D positions, and fuzzy gold standard local scales, respectively. An expert constructs personalized fuzzy gold standard 3D models of the left femur, (M_{FGS}), corresponding to 85 patients, using the semi-automatic commercial software, SterEOS (Chaibi et al., 2012) as an state-of-the-art (SOTA) method. For each patient, the set of 17 fuzzy gold standard handles, $\boldsymbol{CP}^{FGS}$, the corresponding 3D positions, and local scales are extracted from the personalized fuzzy gold standard (M_{FGS}).

5.1.2.4 Gold standard personalized 3D models

To evaluate the accuracy of the 3D reconstructed models by the proposed 3D/2D registration method, we use gold standard personalized 3D models as ground truths. Computerized tomography (CT)-scan-based reconstructed 3D models of the left femur from five unseen pathological patients, who are different from the set of 85 patients, are used as gold standard personalized 3D models, (M_{GS}). CT-scan-based personalized 3D models are reconstructed from CT-scan slices via the SliceOmatic© software (https://www.tomovision.com/products/sliceomatic.html).

5.1.2.5 Training data for regression models

For each CNN_i^{Shape}, corresponding to each handle CP_i^T, we extract 140 local patches as (LIs) (including frontal and lateral), which are described in section 4.4.2.1, from a total of 70 patients' bi-planar radiographs. For all $CNNs^{Shape}$, corresponding to the set of 17 3D handles, we end up with (17×140) (LIs). To train $CNNs^{Scale}$ regression models, the same 70 patients' 2D bi-planar radiographs are considered as reference images. To train each CNN_i^{Scale} corresponding to $CP_i^{\acute{}}$, 140 local patches (LIRs), described in section 4.4.3.1, of 70 patients are computed, leading to (17×140) LIRs for the set of 17 $\boldsymbol{CP}^{FGS}$ to train 17 $CNNs^{Scale}$ models.

5.1.3 Validation protocol for 3D positions and 3D scale ratios of handles

5.1.3.1 Evaluation data and implementation

To evaluate the accuracy of the 3D position and local scale of the 17 handles of the personalized left femur, we automatically estimate 3D displacements of the set of 17 3D handles via 17 trained $CNNs^{Shape}$ models. Then, the 3D scale ratios of the 17 3D handles are automatically estimated via 17 trained $CNNs^{Scale}$ models. The proposed method is implemented in a home-made software application and run by a GeForce® GTX GPU. The CNN-based regression models (U_{Shape}) are developed in a Tensor-flow platform and implemented by a GeForce® GTX 1060 GPU. In each training iteration (epoch = 100), the mini-batch size is equal to 10, and the learning rate is 0.009. The learning rate is a hyper parameter that we can tune it in training process of the CNN. It controls how quickly the model is adapted to the problem. There is a tradeoff between the choice of learning rate and epoch numbers. For example, smaller learning rate needs more training epochs to converge. In non-convex optimization, using mini batch helps to avoid getting trapped in local optimums. Using mini batch updates the model parameters more frequently. Smaller batch size offers a regularization effect and lower generalization error. They are the hyper parameters that we tune in training process of the CNN-based regression models.

5.1.3.2 Evaluation metrics for 3D positions of handles

To evaluate the accuracy of the 3D position of 17 handles, we compute the mean and standard deviation (Mean $\pm$ STD), the maximum (Max), and minimum (Min) of the absolute 3D Euclidean distance error (mm) between displaced 3D positions of the 17 3D handles, $CP^{'}$, computed by the proposed method, and the corresponding 17 fuzzy gold standard 3D positions, CP^{FGS}, over 15 evaluation cases.

5.1.3.3 Evaluation metrics for 3D scale ratios of handles

To evaluate the accuracy of the local scale ratios of 17 handles, we compute the mean and standard deviation (Mean $\pm$ STD), the maximum (Max), and minimum (Min) of the absolute scale ratio errors between the estimated$\left(\delta s_x, \delta s_y, \delta s_z\right)$ and the fuzzy gold standard scale ratios, $\left(\delta s_x, \delta s_y, \delta s_z\right)^{FGS}$ for the 17 handles over 15 evaluation cases. Of the 17 handles $\boldsymbol{CP}^{'}$, $CP_1^{'}$ to $CP_{12}^{'}$ are the point handles with uniform scale $\left(s_x = s_y = s_z\right)_i$, and the other five are the spline handles with non-uniform scale.

5.1.4 Validation protocol for 3D femur and clinical measurements

5.1.4.1 Evaluation data and implementation

To evaluate the proposed fully automatic 3D/2D registration in the personalized 3D femur reconstruction application, we evaluate the accuracy of the personalized 3D shape reconstruction of the left femur. We compare the accuracy of $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ with two different evaluation sets:

1) Fuzzy gold standard (M_{FGS}) of the 15 unseen patients described in section 5.2.2.1;
2) Ground truth gold standard (CT)-scan-based reconstructed 3D shape femurs (M_{GS}) of a new set of five unseen patients described in section 5.2.1.4).

The first evaluation compares our proposed fully automatic method with a semi-automatic approach integrated in a commercial tool, SterEOS (Chaibi et al., 2012), which we use to generate (M_{FGS}) of the same 15 unseen patients. In the second evaluation, we compare the personalized 3D femur $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ with (M_{GS}). We firstly apply the Iterative Closest Point (ICP) rigid registration method (Besl, P. and McKay, 1992) to align the estimated $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ into (M_{GS}), for transfer into the same 3D coordinate system. Then, we compute the 3D shape accuracy between $\left(M_g^{\hat{S}\circ(\widehat{D\circ}(T))}\right)$ and (M_{GS}). In addition to the accuracy of the computation of the personalized 3D shape reconstructed femur, four important clinical 3D measurements of the personalized 3D shape of femurs are computed via a commercial tool integrated in SterEOS software (Chaibi et al., 2012). These extracted 3D measurements are:

1) HKS (the hip knee center-femoral shaft angle, which is the angle between the mechanical and anatomical femoral axis);
2) FMA (the femoral mechanical angle);
3) FT (the femoral torsion, which is the angle between the femoral neck axis and the bi-condoler femoral axis);
4) FL (the femoral length).

5.1.4.2 Evaluation metrics

The accuracy of the reconstructed personalized 3D shape of the left femurs is evaluated based on two different measurements, namely, the point to surface (P2S) distance and the 3D Hausdorff distance errors between the reconstructed 3D femur $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ and the fuzzy gold standard (M_{FGS}) and gold standard (M_{GS}) models. We compute the RMS, Mean, Min, and Max of (P2S) distance errors (mm) between the estimated $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ and the corresponding (M_{FGS}) and (M_{GS}). Then, we compute the 3D Hausdorff distance errors between the estimated $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ and (M_{FGS}). To evaluate clinical the 3D measurements, we compute the Mean and Standard Deviation of Absolute Errors $(\text{MAE} \pm \text{STD})$ of four clinical 3D measurements of the left femur, separately, between the estimated $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ and fuzzy gold standard (M_{FGS}), in mm and degrees. Distance map (colorful-coded error distribution) to evaluate the accuracy of the 2D-biplanar radiographs-based patient-specific 3D femur in comparison with the fuzzy gold standard 3D models (M_{FGS}) for two validation cases with minimum and maximum RMS-P2S errors.

5.2 Results

5.2.1 Results of 3D/2D similarity registration

Tables 5.1 and 5.2 summarize the quantitative results of the three different validation experiments over the evaluation cases for the coarse registration and after the fine registration step (including both coarse and fine registration), respectively. In both coarse and fine registrations, Mean and Standard Deviation of Absolute Pose Errors $(\text{MAE} \pm \text{STD})$ between estimated and fuzzy gold standard 3D bone pose parameters and isotropic scaling are computed over evaluation cases. Tables 5.1 and 5.2 show the comparison of our work with the method in (Chaibi et al., 2012), as an state-of-the-art (SOTA) method, on the same data set. (Chaibi et al., 2012) propose a semi-automatic method with manual 3D pose initialization integrated in the commercial SterEOS software, that we use to generate the fuzzy gold standard 3D models (M_{FGS}).

Table 5.1 (MAE±STD) of estimated 3D pose and isotropic scaling in coarse registration

	(M_{FGS}) into $(Mask_{F,L}^{FGS})$	(M_g) into $(Mask_{F,L}^{FGS})$	***(M_g) into $(Mask_{F,L})$***
3D Errors	PCA +$CNN^1_{(r_\beta)}$	PCA +$CNN^1_{(r_\beta)}$	***PCA +$CNN^1_{(r_\beta)}$***
$t_x(mm)$	4.19 ± 3.95	4.69±2.48	***5.20±2.25***
$t_y(mm)$	2.00±1.22	2.10±1.42	***2.84±1.84***
$t_z(mm)$	1.76±1.42	1.91±1.12	***2.56±1.48***
$r_\theta(°)$	3.46±0.92	3.52±1.29	***3.96±1.37***
$r_\alpha(°)$	3.84±1.92	4.02±1.39	***4.85±1.28***
$r_\beta(°)$	3.93±2.93	4.53±3.27	***4.95±3.68***
s(%)	0.01±0.00	0.04±0.03	***0.05±0.04***

Table 5.2 (MAE±STD) of estimated 3D pose in fine registration step

	(M_{FGS}) into $(Mask_{F,L}^{FGS})$	(M_g) into $(Mask_{F,L}^{FGS})$	***(M_g) into $(Mask_{F,L})$***
3D Errors	fine CNNs	fine CNNs	***fine CNNs***
$t_x(mm)$	0.12±0.09	0.14±0.12	***0.27±0.29***
$t_y(mm)$	0.14±0.10	0.15±0.11	***0.16±0.12***
$t_z(mm)$	0.13±0.11	0.16±0.10	***0.16±0.12***
$r_\theta(°)$	0.27±0.12	0.29±0.11	***0.30±0.07***
$r_\alpha(°)$	0.32±0.14	0.34±0.13	***0.37±0.16***
$r_\beta(°)$	0.23±0.22	0.29±0.29	***0.33±0.26***

5.2.2 Results of 3D/2D non-rigid registration

5.2.2.1 3D position accuracy of handles

Figure 5.1 illustrates the quantitative results of the computed 3D position accuracy between displaced 3D positions, $CP^{'}$, computed by the proposed method, and the corresponding fuzzy gold standard 3D positions, CP^{FGS}, of 17 3D handles. For each handle, the black dot, red line, top and bottom of blue line caps show $(\text{Mean} \pm \text{STD})$, (Max), and (Min) errors, respectively.

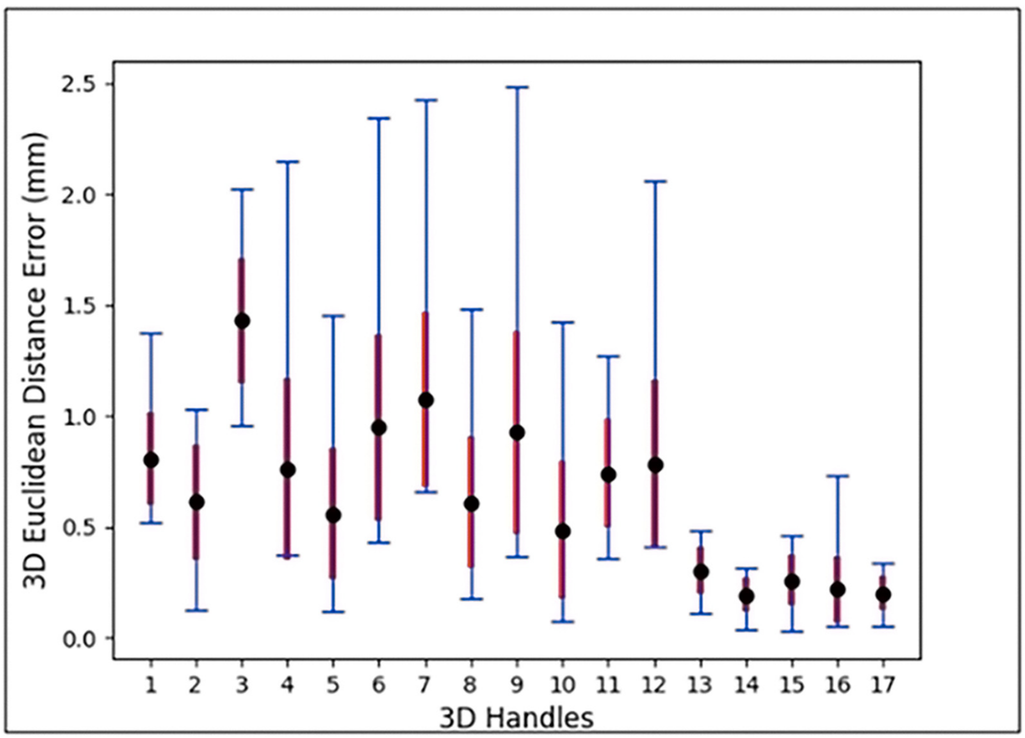


Figure 5.1 Error bars of computed 3D positions of 17 3D handles

5.2.2.2 3D scale ratio accuracy of handles

Figure 5.2 illustrates the quantitative results of the local 3D scale ratio accuracy between the estimated $(\delta s_x, \delta s_y, \delta s_z)$, via the CNN^{Scale} regression models, and the fuzzy gold standard scale ratios, $(\delta s_x, \delta s_y, \delta s_z)^{FGS}$ for 17 3D handles. For each handle, the black dot, red line, top and bottom of blue line caps show the (Mean $\pm$ STD), (Max), and (Min) errors, respectively. In Figure 4.17, for 5 non-uniform handles, CP'_{13} to CP'_{17}, we show the maximum scale errors among the X-, Y-, and Z-axes.

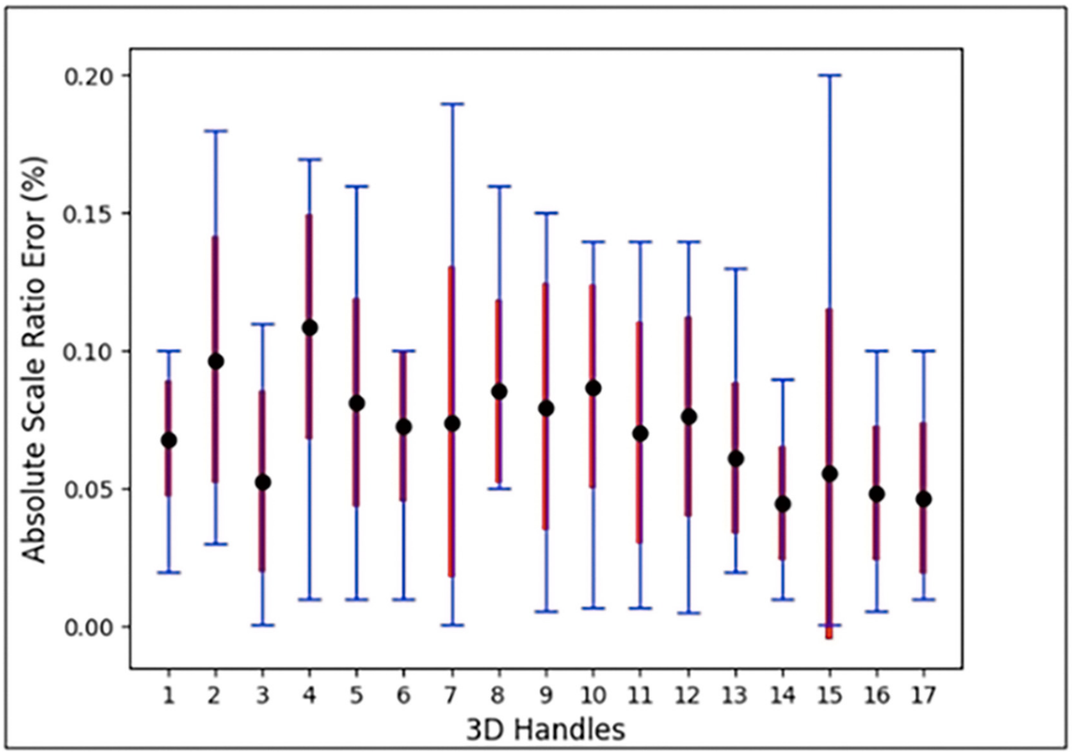


Figure 5.2 Error bars of local 3D scale ratios of 17 3D handles

5.2.2.3 Personalized 3D femur accuracy

Table 5.3 presents the average of (RMS, Mean, STD, Min, Max) of (P2S) distance errors (mm) between personalized 3D femurs $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ and the corresponding fuzzy gold standards (M_{FGS}) over 15 evaluation cases. Table 5.3 shows the comparison of our work with the method in (Chaibi et al., 2012), as an state-of-the-art (SOTA) method, on the same data set. (Chaibi et al., 2012) propose a semi-automatic method integrated in the commercial SterEOS software, that we use to generate the fuzzy gold standard 3D models (M_{FGS}). The

$(\text{Mean} \pm \text{STD})$ of the 3D Hausdorff distance errors over 15 unseen patients is equal to (2.95$\pm$1.42) mm. Table 5.4 presents the mean and standard deviation of absolute errors of four clinical 3D measurements extracted from personalized 3D femurs in comparison with the corresponding fuzzy gold standard models of 15 unseen patients. Table 5.4 shows the comparison of the clinical 3D parameter accuracy with the method in (Chaibi et al., 2012), as an state-of-the-art method, on the same data set. Table 5.5 for its part presents the accuracy of the personalized 3D femurs $\left(M_g^{\hat{S}\circ(\hat{D}\circ(T))}\right)$ in comparison with the ground truth gold standard 3D models (M_{GS}). The average of (RMS, Mean, STD, Min, Max) of point to surface (P2S) distance errors (mm) are computed over five unseen patients.

Table 5.3 Average of (RMS, Mean, STD, Min, Max) of P2S errors
in comparison to fuzzy gold standard 3D models by (SOTA)

	RMS (mm)	Mean (mm)	STD (mm)	Min (mm)	Max (mm)
Average	0.88	0.66	0.57	9.77E-05	2.87

Table 5.4 (MAE ±STD) of four clinical 3D measurements in degrees and (mm)
on 15 reconstructed 3D femurs comparing to fuzzy gold standard models by (SOTA)

	HKS(°)	FMA(°)	FT (°)	FL (mm)
MAE	0.14	0.09	0.16	0.67
STD	0.09	0.09	0.12	0.35

Table 5.5 Average of (RMS, Mean, STD, Min, Max) of P2S distance error
as compared to CT-scan-based gold standard 3D models on 5 validation cases

	RMS (mm)	Mean (mm)	STD (mm)	Min (mm)	Max (mm)
Average	2.70	2.14	1.64	0	11.08

Figures 5.3 and 5.4 show superposition of the 2D bi-planar radiographs-based reconstructed 3D femurs (orange) and the corresponding fuzzy gold standard 3D models (red) on patients' 2D bi-planar radiographs in $45^{\circ}/45^{\circ}$ and $0^{\circ}/90^{\circ}$ orientations, respectively. In addition, Figures 5.5 and 5.6 illustrate color distance map bar between 2D-bi-planar based reconstructed 3D femur and fuzzy gold standard 3D femur (M_{FGS}) of above mentioned 3D femurs corresponding to Figure 5.3 and 5.4, respectively.

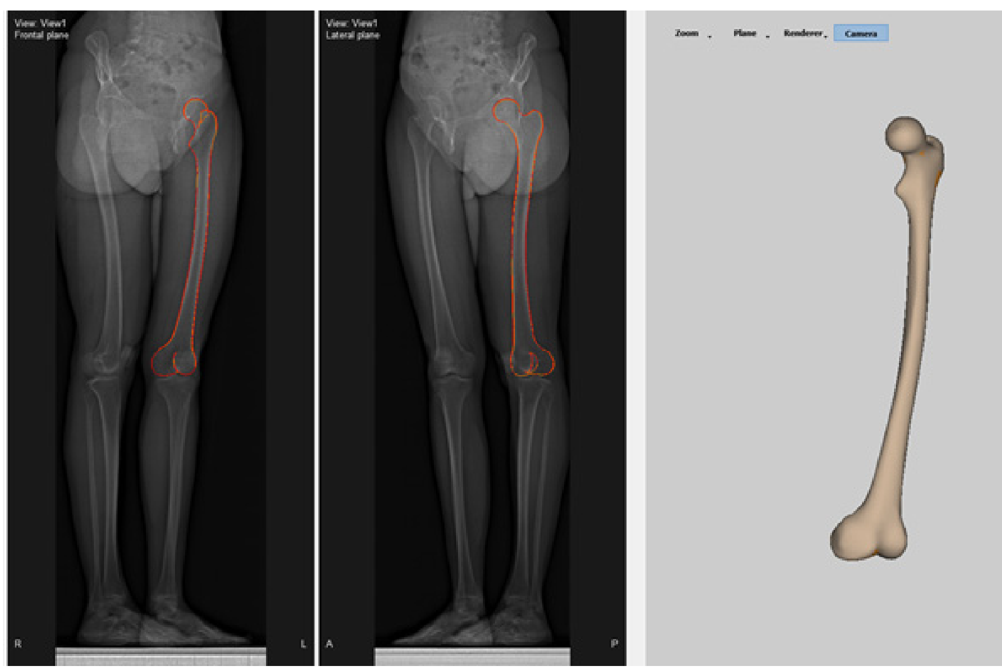


Figure 5.3 (Left and Middle) 2D projection of reconstructed 3D femur (orange) and fuzzy gold standard (red) on radiographs, (Right) reconstructed 3D femur

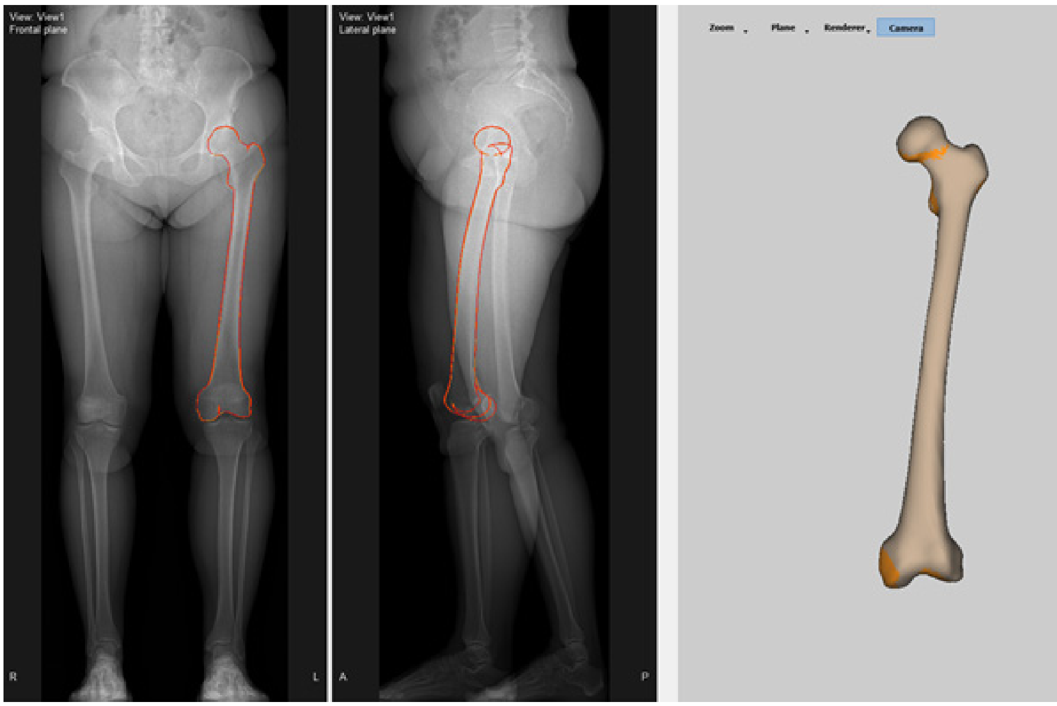


Figure 5.4 (Left and Middle) 2D projection of reconstructed femur (orange) and fuzzy gold standard (red) on radiographs, (Right) reconstructed 3D femur

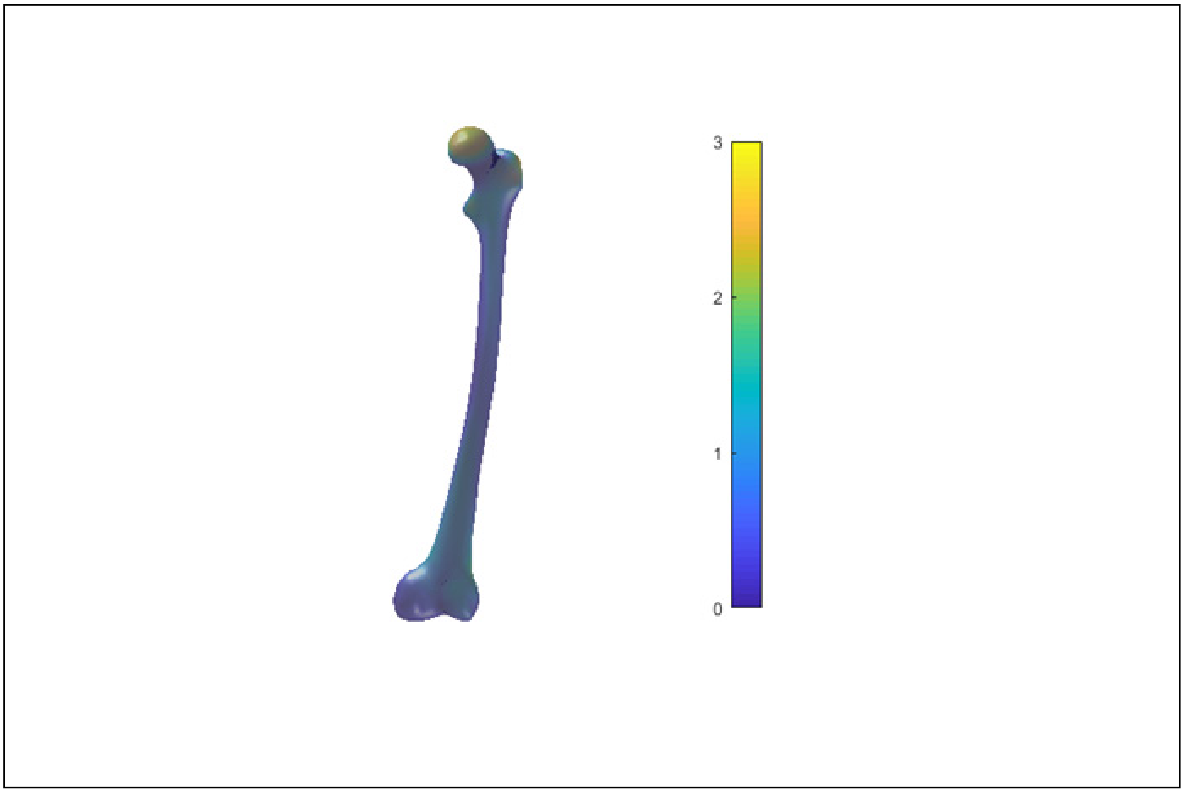


Figure 5.5 Illustrates color distance map bar between 2D-bi-planar based reconstructed 3D femur with RMS-P2S error of 0.66 mm

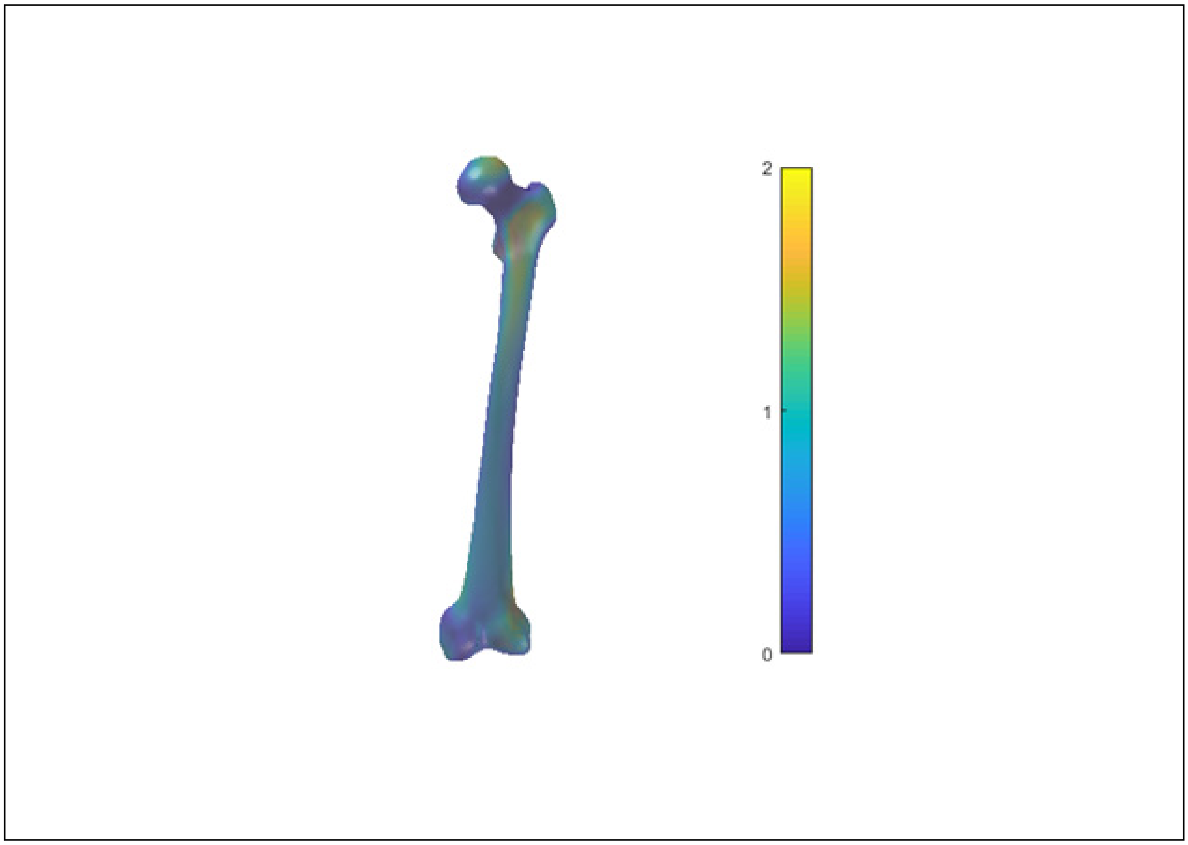


Figure 5.6 Illustrates color distance map bar between 2D-bi-planar based reconstructed 3D femur with RMS-P2S error of 0.5 mm

CHAPITRE 6

DISCUSSION

6.1 Discussion automatic 3D femur pose and isotropic scale estimation

The performance and accuracy of the 3D femur pose and isotropic scale estimation is evaluated in three different experiments. The first experiment conducts a 3D/2D similarity registration from fuzzy gold standard 3D models (M_{FGS}), as a simulated model, into the corresponding fuzzy gold standard masks $(Mask_F^{FGS}, Mask_L^{FGS})$, as simulated masks. The second experiment registers $\left(M_g\right)$ into $(Mask_F^{FGS}, Mask_L^{FGS})$, to validate the effects of the generic 3D model as a source. Ultimately, the third experiment investigates a real application; 3D/2D similarity registration of $\left(M_g\right)$ into CNN-based $(Mask_F, Mask_L)$.

In Table 5.1, the first and second experiments, show using the generic 3D model of the femur $\left(M_g\right)$ rather than fuzzy gold standard 3D model of the femur (M_{FGS}) has a minor effect on the 3D pose estimation accuracy. The comparison between the first and second experiment validates the feasibility of the proposed methodology without possible errors introduced by the generic 3D model and the CNN-based segmented masks. The comparison between the second and third experiments demonstrates that results of the coarse registration are significantly sensitive to the accuracy of the segmented bone structure. Particularly, superposition of the left and right femur in the lateral projection of the patients, in $0^{\circ}/90^{\circ}$ orientation affects the accuracy of (t_x). However, Table 5.2 shows that the refine CNNs-based regression models noticeably improve the accuracy of all six 3D pose parameters to sub-millimeter and sub-degree precision. In the fine registration step, the isotropic scaling ratio(s) is not refined (Table 5.2), since in the coarse registration we achieve sub-percentage precision (Table 5.1).

Despite difficulties to compare the performance of different approaches owing to different databases, experiments, and applications (Hatt et al., 2015 ; Miao et al., 2016 ; Zheng et al., 2018), among the most relevant results in orthopedic applications, (Baka et al., 2011) report absolute 3D pose errors of the similarity registration of 3D statistical model of the distal femur into the 2D stereo Radiographs. In comparison with (Baka et al., 2011), the mean of (MAE) for (t_x, t_y, t_z),$(r_\theta, r_\alpha, r_\beta)$, and (s) are reduced from 0.86 mm, 1.30°, and 1.55(%) to 0.19 mm, 0.33°, and 0.05(%), respectively. However, Table 5.1 and Table 5.2 show the comparison of our work with (Chaibi et al., 2012) for the same data set. (Chaibi et al., 2012) propose a semi-automatic method integrated in the commercial SterEOS software, that we use to generate the fuzzy gold standard 3D models (M_{FGS}).

Our proposed fully automatic approach achieves (MAE) of proximal distal axis rotation, which is the most difficult parameter to estimate (Baka et al., 2011 ; Miao et al., 2016 ; Zheng et al., 2018). The proposed fully automatic coarse-to-fine CNNs-based 3D/2D registration approach achieves highly accurate and robust 7DOF 3D pose and isotropic scaling to automatically fit a generic 3D model to 2D bi-planar radiographs. The success rate is of 100(%) at MAE lower than 1 mm, 1°, and 0.1(%). The developed system could be feasible to apply for other lower limb 3D bone structures such as the tibia. In addition, the proposed fully automatic coarse-to-fine CNNs-based 3D/2D registration could be tested to apply on the patient's radiographs with knee flexions, since T, in the fine step $CNN^3_{(\delta t_x, \delta r_\alpha)}$, which is responsible to estimate out-of-plane rotation residuals,δr_α, is trained to cover the range of $[-15, +15]$.

6.2 Discussion automatic 3D femur reconstruction

The ultimate goal is to have a personalized 3D femur reconstruction and clinical 3D measurements from 2D bi-planar EOS® radiographs. In 3D/2D non-rigid registration, we successfully merge CNN cascade-based regression models with MLS 3D shape and scale deformation (Table 5.3). Unlike a semi-automatic 3D femur reconstruction method integrated in SterEOS (Chaibi et al., 2012), which needs user interventions to annotate the new positions of 17 3D control and manually adjust the scales, our CNN-based regressors provide significant advantages for:

1) Removing operator interventions;
2) Fast adjustments of 3D handles' positions and scales, to drive MLS 3D shape and scale deformation (Figure 5.1).

In comparison with fuzzy gold standard 3D handle positions, the average of Mean Absolute 3D Euclidean distance Errors (MAE) of the 17 3D handles on 15 validation cases is equal to 0.63 mm. Using CNN allows to achieve the Mean of 3D Euclidean distance errors on 15 validation cases lower than 1 mm for the set of 3D handles, rather than C3 (1.42 mm) and C7 (1.07 mm) (Figure 5.1). C3 and C7 correspond to a lesser trochanter and condyle post exterior, respectively, which are less visible than other 3D handles because of the overlapping of the right and left femur in patients with a $0^{\circ}/90^{\circ}$ and $45^{\circ}/45^{\circ}$ orientation, respectively. Unlike (Yu et al., 2016), which uses FFD deformation on a large set of 3D control points (88), having a small set of (17) handles helps us apply MLS deformation, which allows clinicians to easily and manually correct the reconstructed 3D femur errors. (Yu et al., 2016) uses the triangulation-based method to compute the corresponding 3D positions of updated 2D bi-planar positions of the projected 3D control points; this is done by direct estimation of the 3D displacements at the origin $(0,0,0)$ of the 3D referential of the EOS® cabin system. By contrast, our contribution to a 3D/2D non-rigid registration involves fewer steps.

The 17 CNN-based local 3D scale ratio estimations developed automatically provides personalized local scale ratios corresponding to 17 3D handles. Merging CNN-based local 3D scale ratio regressors with MLS deformation automatically adjusts the 2D silhouette of the 3D femur with the edges of the target bone in patients' 2D bi-planar radiographs. Compared to fuzzy gold standard 3D handles' scale ratios, using CNN to estimate the local scale ratios of 17 3D handles provides an (MAE) lower than $0.1(\%)$ over 15 validation cases for each handle (Figure 5.2).The mean of the (MAE) of the 17 3D handles is equal to $0.05(\%)$. The Min $\left(Es_x=0.05, Es_y=0.03, Es_z=0.05\right)$ and Max (0.1) of (MAE) are C15 (spline diaphysis with non-uniform scale) and C7 (condyle post exterior with uniform scale), respectively.

The ultimate personalized 3D femur accuracy is validated in comparison with two different validation sets. In the first validation, we compare the accuracy of the personalized 3D femurs with fuzzy gold standard 3D models reconstructed via the semi-automatic approach integrated in the commercial software, SterEOS (Chaibi et al., 2012). The $(\text{Mean} \pm \text{STD})$ of RMS-P2S errors over 15 validation cases are equal to (0.88±0.29) mm (Table 5.3). The (Mean) of Max-P2S errors is (2.87) mm (Table 5.3). For the second validation set, the $(\text{Mean} \pm \text{STD})$ of RMS-P2S errors of five personalized 3D femurs in comparison with CT-scan-based ground truth 3D models are equal to (2.70± 0.39) mm (Table 5.5). The (Mean) of Max-P2S errors is (11.08) mm (Table 5.5). The $(\text{Mean} \pm \text{STD})$ of RMS-P2S errors on the first validation set (Fuzzy gold standard 3D models), (0.88±0.29) mm, is lower than the $(\text{Mean} \pm \text{STD})$ of RMS-P2S errors on the second validation set (CT-scan-based gold standard 3D models), (2.70± 0.39) mm. The differences between these two evaluation metrics results arise from the fact that, the reference 3D models in the second validation set, are the ground truths CT-scan-based 3D models which are reconstructed by manual segmentations of the femur bone boundaries on a stack of the CT slices, So, the ground truth CT-scans-based 3D models define an accurate shape of the femur. In contrast, the reference 3D models in the first validation set, are the fuzzy gold standards 3D models constructed by an expert via a semi-automatic method integrated in the commercial software tool (SterEOS), in which a generic 3D model of the femur is deformed by manual adjustments of a set of the handles to fit into the femur bone structure in 2D radiographs. The reference ground truths CT-scans-based 3D shape of the femur is more accurate than the reference fuzzy gold standard 3D shape. Therefore, the 3D shape accuracy of the personalized 3D femurs are not the same on two different validation sets (Table 5.3 and Table 5.5).

Notwithstanding difficulties in comparing the accuracy of different methods validated based on non-similar databases, experiments, evaluation metrics, and applications, we compare the accuracy of the personalized 3D femurs with the most relevant state-of-the art results in 3D model-based orthopedic applications. Furthermore, the performance of the proposed 3D femur reconstruction method is compared with (Chaibi et al., 2012) on the same data set. In contrast to (Cresson et al., 2010), a semi-automatic 3D femur reconstruction with a manual

3D pose initialization, which optimize MLS handles of the generic 3D femur via a conventional iterative method, we train CNN cascade-based regression models to optimize the MLS handles of the generic 3D femur. In (Cresson et al., 2010), the average of Mean-P2S distance errors on six femurs is 1.0 mm, and the Max of Mean-P2S is 5.53 mm. In contrast to conventional iterative optimization methods in semi-automatic 3D/2D non-rigid SSM registration (Baka et al., 2011, 2012 ; Youn et al., 2017 ; Yu et al., 2017, 2015), we train CNN cascade-based regression models for an automatic 3D/2D non-rigid registration of the generic 3D model of the femur in patients' 2D bi-planar EOS® radiographs, in a time-efficient manner. In (Yu et al., 2016), the proposed automatic 3D/2D registration uses FFD deformation with a large set of control points which needs to define a strong regularization term to avoid undesirable distortion in 3D bone reconstruction. Moreover, after the 3D reconstruction process, if a manual adjustment were needed to correct 3D model errors, it would be difficult for clinicians to manually adjust a large set of control points. Hence, our proposed automatic CNN cascade-based 3D/2D registration is more appropriate for clinicians in clinical routine thanks to the use of the MLS deformation, which is deployed as part of the commercial SterEOS software tools (Chaibi et al., 2012) and provides an as-rigid-as-possible shape deformation without undesirable distortion on a small set of 3D handles.

Table 5.3 and Table 5.4 compare the performance of the proposed 3D femur reconstruction method with (Chaibi et al., 2012) for the same data set. Tables 5.3 and 5.4 illustrate the accuracy of the personalized 3D femurs and clinical 3D measurements, respectively, via our automatic CNN cascade-based 3D/2D registration, as compared to the MLS deformation integrated in the semi-automatic commercial SterEOS tools (Chaibi et al., 2012). This is done on the same 15 validation cases. After 3D femur reconstruction, we compute four clinical measurement errors for comparison with the semi-automatic software SterEOS tools (Chaibi et al., 2012), on the same validation set. The MAE and STD of each clinical 3D measurement are lower than 1 mm or 1 degree (Table 5.4). In the clinical routine of 3D model-based orthopedic applications, the proposed automatic and fast CNN cascade-based 3D/2D registration achieves a high accuracy for clinical 3D measurement computation to assist clinicians quickly diagnose and analyze 3D shape deformities in the femur. We achieve

100% success rate for RMS-P2S errors lower than 1 mm versus fuzzy gold standard models (Table 5.3).

The computation time for the 3D femur reconstruction via the proposed fully automatic 3D/2D registration framework including 3D/2D similarity registration and 3D/2D non-rigid registration stages, without any code optimization, is 75 (s), using an Intel® Core 17 CPU and a GeForce® GTX 1060 GPU. Using CNN-based regressors helps to improve the computation time as compared to (Chaibi et al., 2012), which is a semi-automatic method and requires 10 (min) for the reconstruction of both lower limbs with CPU computation. However, in (Cresson et al., 2010), the semi-automatic 3D femur reconstruction method takes around 30 (s) with CPU computation. In (Yu et al., 2015), a conventional iterative optimization is used to reconstruct a personalized 3D proximal femur, with a reported time of 15 (min), with CPU computation. However, in (Yu et al., 2017), the computation time of (Yu et al., 2015) is improved to 1.09 (s) with GPU computation.

The presented fully automatic cascade CNN-based 3D femur reconstruction from 2D bi-planar radiographs efficiently fits the generic 3D model of the left femur into the EOS® 2D bi-planar radiographs acquired with two different fields of view, full body and whole lower limbs, and patients' orientations in $0^{\circ}/90^{\circ}$and $45^{\circ}/45^{\circ}$, in a time efficient manner. The developed fully automatic 3D femur reconstruction system is feasible to be applied for other lower limb 3D bones such as the tibia. As limitations of the proposed method, firstly the 3D pose and isotropic scaling in the coarse step of the 3D/2D similarity registration depends on the accuracy of the segmented mask. However, we refine the 3D pose parameters in the fine registration step. Secondly, the training data generation separately for two main stages is time consuming. Thirdly, in the case of having 2D radiographs with marginal spacing problem that some part of the left femur, particularly around the greater trochanter is out of marginal spacing bound of 2D radiographs, we could not adjust 3D handle's position and local scale corresponding to the greater trochanter.

Tables 6.1 illustrates the accuracy of the personalized 3D femurs in comparison to state-of-the art methods, while using CT-scan-based reconstructed 3D model as ground truth model. In addition, Table 6.1 shows the 3D femur reconstruction time of the proposed methodology in comparison with the previous state-of-the-art methods.

Table 6.1 Average of P2S errors of the reconstructed 3D femurs via our proposed method comparing to previous state-of-the-art methods

Authors & Date	3D Reconstruction method	Validation Modality	3D Reconstruction Accuracy (P2S) (mm) and Time
(Laporte et al.,2003)	Manual initialization With Generic model for 3D distal femur reconstruction	CT	Mean = 1.0 RMS = 2.8 Time = 15 minutes
(Cresson et al., 2010)	Semi-automatic with MLS deformation for 3D femur reconstruction	CT	Mean = 1.0 Time = 1 minute
(Zheng G., 2011)	Manual initialization with SSM for 3D proximal femur reconstruction	CT	Mean = 1.5 Time = 15 minutes
(Baka et al., 2011)	Manual initialization with SSM for 3D distal femur reconstruction	CT	RMS = 1.68 Time = 5 minutes
(Chaibi et al., 2012)	Semi-automatic with MLS deformation for whole lowe limbs reconstruction	CT	Mean = 1.2 RMS = 3.2 Time = 5 minutes for one lower limbs
(Yu et al., 2016)	Automatic initialization Free From Deformation for proximal femur reconstruction	CT	Mean = 1.3 Time = Not reported
(Yu et al., 2017)	Manual Initialization Free Form Deformation for proximal femur reconstruction	CT	Mean = 1.2 Time = 1.09 seconds
Our methodology	Automatic initialization with CNN-based MLS deformation for 3D femur reconstruction	CT	Mean = 2.14 RMS = 2.70 Time = 75 seconds

CONCLUSION

In this thesis, we presented a fully automatic cascade CNN-based lower limb bone 3D reconstruction framework via two main stages: 3D/2D similarity registration and 3D/2D non-rigid registration. In the first stage of the lower limb bone 3D reconstruction, we presented a fully automatic coarse-to-fine CNN-based registration approach. In the automatic 3D/2D similarity registration step, we achieved highly accurate and robust 7DOF3D pose and isotropic scaling to automatically fit a generic 3D model of the left femur into the EOS®2D bi-planar X-ray images. In the second stage, we presented a fully automatic CNN cascade-based 3D/2D non-rigid registration framework to efficiently deform the local 3D shape and 3D scale of the 3D femur. The proposed CNN-based 3D handle displacement and scale ratio estimation eliminates manual annotations and user interventions for MLS deformation, and does so in a time-efficient manner. This method provides for physicians a capacity, if required, to easily and manually adjust possible errors of the reconstructed 3D model. We achieved an average of RMS-P2S accuracy of (0.88) and (2.70) mm in evaluation by fuzzy gold standard and CT-scan-based ground truth gold standard 3D models, respectively. When compared to related works, the results obtained with the proposed system are close, and the system could be used efficiently for other bone structures, such as the tibia.

7.1 Publications and presentations

7.1.1 Peer review journal

Babazadeh Khameneh, N., Vazquez, C., Cresson, T., Lavoie, F., & de Guise, J. (2022). Automatic CNN-based 3D/2D non-rigid registration method for fast 3D femur reconstruction from bi-planar radiographs. Submitted in Journal of Medical Imaing (JMI), SPIE, December 2022.

7.1.2 International conferences

Babazadeh Khameneh, N., Vazquez, C., Cresson, T., Lavoie, F., & de Guise, J. (2021). Highly accurate automated patient-specific 3D bone pose and scale estimation using bi-planar pose-invariant patches in a CNN-based 3D/2D registration framework. *Proceedings - International Symposium on Biomedical Imaging*, *2021-April*(Isbi), 681‑684. https://doi.org/10.1109/ISBI48211.2021.9433843

Babazadeh Khameneh, N., Cresson, T., Bakhous C., Vazquez, C., & de Guise, J. (2018). Automatic 3D pose estimation via co-trained CNN models. 39th SICOT.

7.1.3 Local Conferences

Babazadeh Khameneh, N., Cresson, T., Bakhous C., Vazquez, C., & de Guise, J. (2018). Multi-trained ensemble CNN regression models for 3D pose estimation. 38th Journée de la Recherche du POES et de la Division d'Orthopédie de l'Université de Montréal, L'Hôpital Sainte-Justine, Montréal, Québec, Canada.

Babazadeh Khameneh, N., Cresson, T., Bakhous C., Vazquez, C., & de Guise, J. (2018). Fully automatic 3D pose estimation. 10th Journée scientifique des etudiantes du Centre de recherche du Centre Hospitalier Université de Montréal (CRCHUM), Montréal, Québec, Canada.

7.2 Recommendation and future work

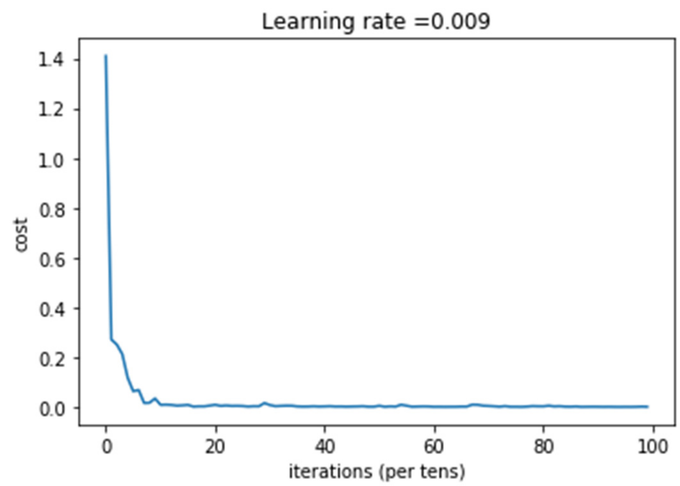
The validation results of the proposed methodology on 3D femur reconstruction are promising and the proposed system could be successfully applied to other bone structures such as the tibia. Using the proposed method for the tibia provides the capability to measure the tibia's clinical 3D geometrical parameters. The context of the 3D femur model reconstruction is rich with potential applications and development of the prototypes opens the door to possible research opportunities. Following the accomplishment of this project, as future work, we recommend to apply the proposed system on 3D pose estimation of the knee flexion, since we trained CNN-based 3D pose estimation models with generated DRRs of the femur with the knee flexions until 15 degrees. It will need some modification on the model's development and then validation of the model, since the structure of the knee bone differs from the femur and we developed this 3D pose estimation model for the femur bone. Moreover, in both 3D pose estimation and 3D shape deformation stages of the project, we can use data augmentation techniques and add more images to training data set to improve the CNN-based regression models accuracy. Particularly, we can retrain, with the augmented data set, two CNNs corresponding to two 3D handles' displacement estimation, C3 (lesser trouchanter) and C7 (condyle post exterior), which are less visible than other 3D handles because of the overlapping of the right and left femur in patients with a $0^\circ/90^\circ$ and $45^\circ/45^\circ$ orientation, respectively.

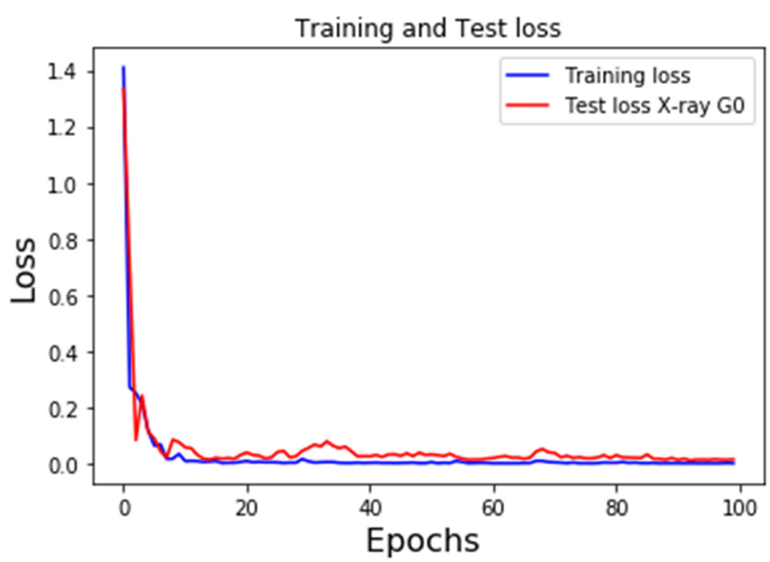
ANNEX I

CONVOLUTIONAL NEURAL NETWORK (CNN)-BASED MODELS PARAMETERS AND PERFORMANCE

CNN-based model performance in the coarse registration step:

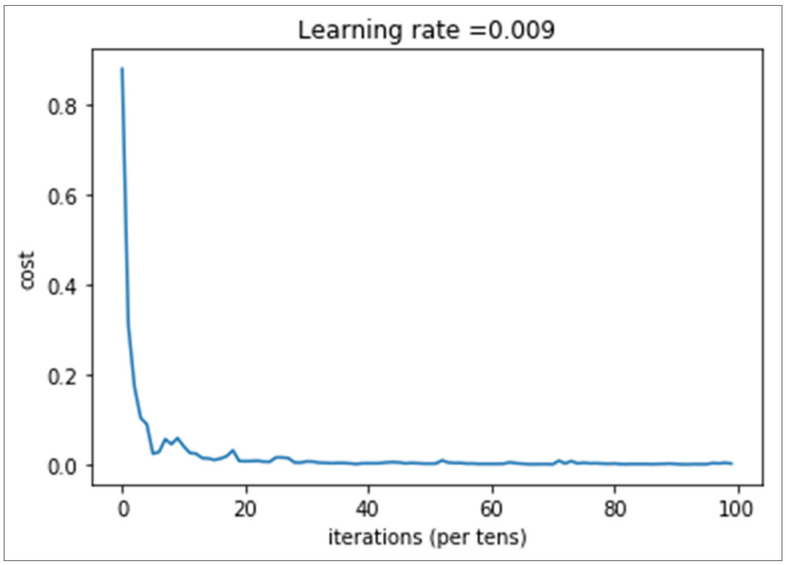
- Epoch Number: 99
- Cost after epoch 99: 0.001000
- Iteration 99, Loss= 0.001000, Training error= [0.0007805552]
- Optimization Finished!

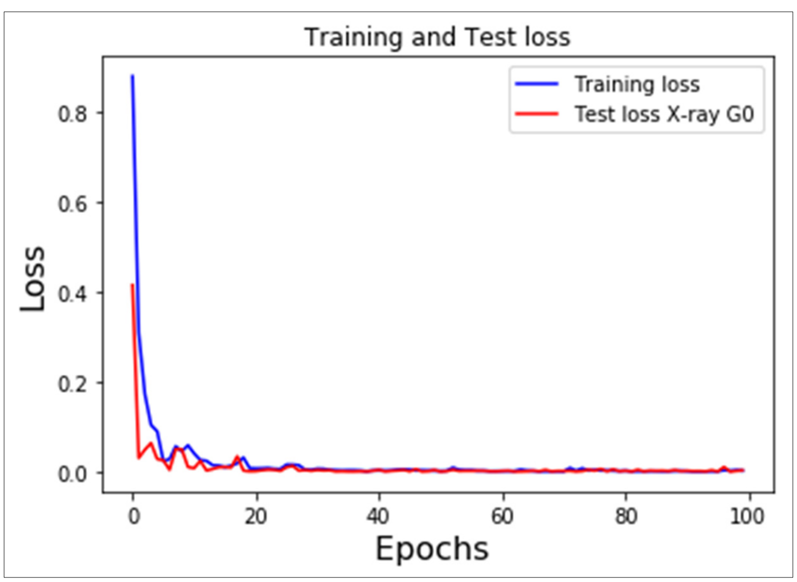




CNN-based model performance in the fine registration step:

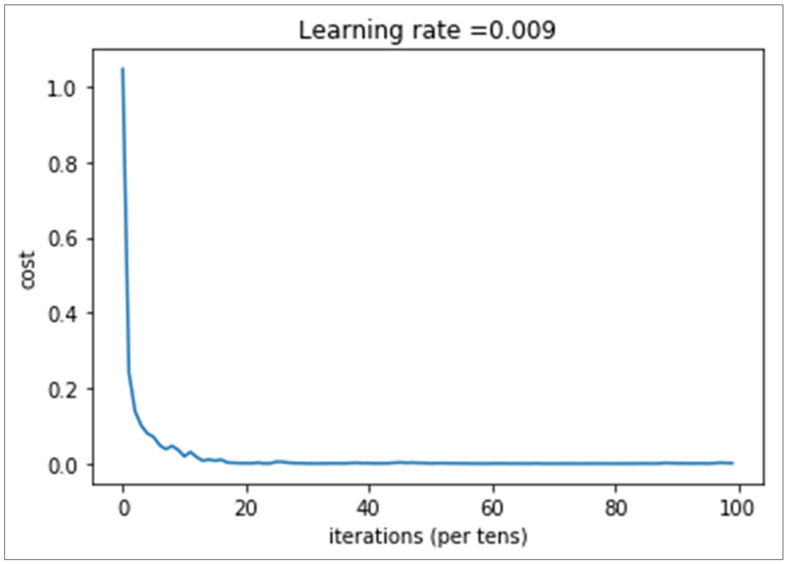
- Epoch Number: 99
- Cost after epoch 99: 0.002088
- Iteration 99, Loss= 0.002088, Training error= [0.0025451826]
- Optimization Finished!

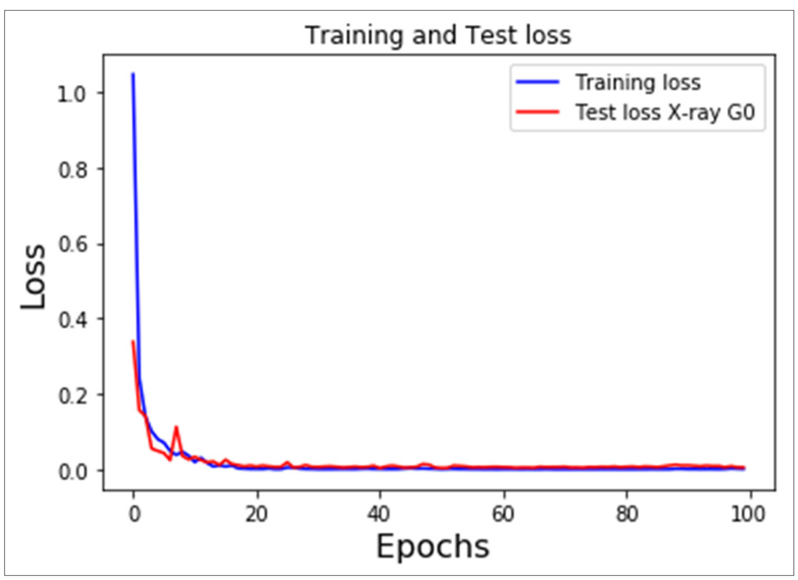




CNN-based model performance in 3D handles displacement estimation:

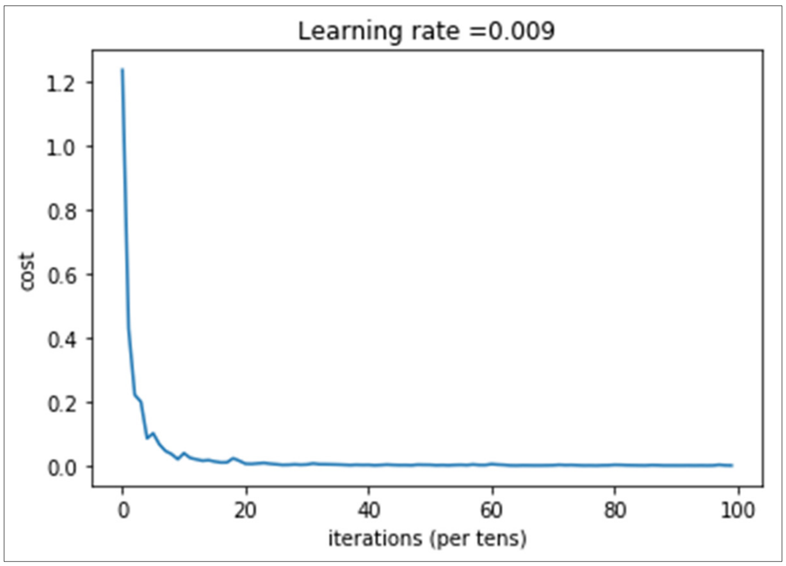
- Epoch Number: 99
- Cost after epoch 99: 0.001232
- Iter 99, Loss= 0.001232, Training error= [0.00037548546]
- Optimization Finished!

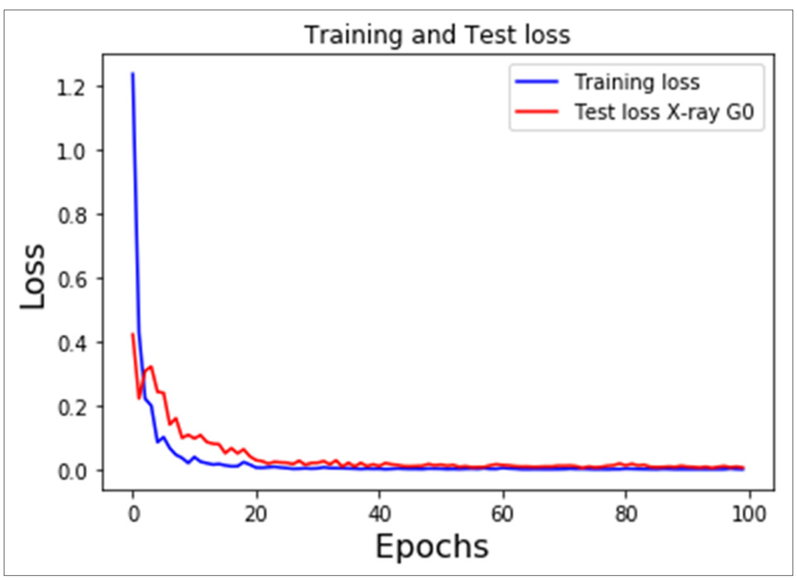




CNN-based model performance in 3D handles scale ratio estimation:

- Epoch Number: 99
- Cost after epoch 99: 0.000969
- Iter 99, Loss= 0.000969, Training error= [0.00080144696]
- Optimization Finished!





LIST OF BIBLIOGRAPHICAL REFERENCES

Abebe, E. S., Utturkar, G. M., Taylor, D. C., Spritzer, C. E., Kim, J. P., Moorman, C. T., … DeFrate, L. E. (2011). The effects of femoral graft placement on in vivo knee kinematics after anterior cruciate ligament reconstruction. *Journal of Biomechanics*, *44*(5), 924‑929. https://doi.org/10.1016/j.jbiomech.2010.11.028

Agomma, R. O., Vazquez, C., Cresson, T., & de Guise, J. (2019). Detection and identification of lower-limb bones in biplanar x-ray images with arbitrary field of view and various patient orientations. *Proceedings - International Symposium on Biomedical Imaging*, *2019-April*(Isbi), 1174‑1177. https://doi.org/10.1109/ISBI.2019.8759456

Anderst, W., Zauel, R., Bishop, J., Demps, E., & Tashman, S. (2009). Validation of three-dimensional model-based tibio-femoral tracking during running. *Medical Engineering and Physics*, *31*(1), 10‑16. https://doi.org/10.1016/j.medengphy.2008.03.003

Arun, K. S., Huang, T. S., & Blostein, S. D. (1987). Least-squares fitting of two 3-D point sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, *9*(5), 698‑700.

Aubert, B., Vazquez, C., Cresson, T., Parent, S., & de Guise, J. (2019). Toward Automated 3D Spine Reconstruction from Biplanar Radiographs Using CNN for Statistical Spine Model Fitting. *IEEE Transactions on Medical Imaging*, *38*(12), 2796-2806. https://doi.org/10.1109/tmi.2019.2914400

Baka, N., de Bruijne, M., van Walsum, T., Kaptein, B. L., Giphart, J. E., Schaap, M., … Lelieveldt, B. P. F. (2012). Statistical Shape Model-Based Femur Kinematics From Biplane Fluoroscopy. *IEEE Transactions on Medical Imaging*, *31*(8), 1573‑1583. https://doi.org/10.1109/TMI.2012.2195783

Baka, N., Kaptein, B. L., de Bruijne, M., van Walsum, T., Giphart, J. E., Niessen, W. J., & Lelieveldt, B. P. F. (2011). 2D-3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. *Medical Image Analysis*. https://doi.org/10.1016/j.media.2011.04.001

Baka, N., Metz, C. T., Schultz, C. J., Van Geuns, R. J., Niessen, W. J., & Van Walsum, T. (2014). Oriented Gaussian mixture models for nonrigid 2D/3D coronary artery registration. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2014.2300117

Besl, P. and McKay, N. (1992). A method for registration of 3-D Shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, *14*, 239‑256.

Cerveri, P., Belfatto, A., & Manzotti, A. (2020). Predicting Knee Joint Instability Using a Tibio-Femoral Statistical Shape Model. *Frontiers in Bioengineering and Biotechnology*, *8*(April), 1‑12. https://doi.org/10.3389/fbioe.2020.00253

Chaibi, Y., Cresson, T., Aubert, B., Hausselle, J., Neyret, P., Hauger, O., … Skalli, W. (2012). 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. https://doi.org/10.1080/10255842.2010.540758

Chen, C., & Zheng, G. (2014). Fully automatic segmentation of AP pelvis X-rays via random forest regression with efficient feature selection and hierarchical sparse shape composition. *Computer Vision and Image Understanding*, *126*, 1‑10. https://doi.org/10.1016/j.cviu.2014.04.015

Cootes, T. F., Taylor, C. J., Cooper, D., & Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*, *61*, 38-59. https://doi.org/10.1006/cviu.1995.1004

Cresson, T., Branchaud, D., Chav, R., Godbout, B., & de Guise, J. (2010). *3D shape reconstruction of bone from two x-ray images using 2D/3D non-rigid registration based on moving least-squares deformation*. *Medical Imaging 2010: Image Processing*. https://doi.org/10.1117/12.844098

Cresson, T., Godbout, B., Branchaud, D., Chav, R., Gravel, P., & de Guise, J. (2008). Surface reconstruction from planar x-ray images using moving least squares. Dans *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. https://doi.org/10.1109/IEMBS.2008.4650078

Cretu, B., Panti, Z., Nica, M., Serban, B., Popa, M., Ene, R., & Cîrstoiu, C. (2018). Lower Limb Deformity and Total Knee Replacement. *Romanian Journal of Orthopaedic Surgery and Traumatology*, *1*(2), 121‑127. https://doi.org/10.2478/rojost-2018-0090

Cuingnet, R., Prevost, R., Lesage, D., Cohen, L. D., Mory, B., & Ardon, R. (2012). Automatic detection and segmentation of kidneys in 3D CT images using random forests BT. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, 66‑74.

Cuno, A., & Esperan, C. (2007). 3D as-rigid-as-possible deformations using MLS 3D As-Rigid-As-Possible Deformations Using MLS, (January).

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology, 26(3), 297-302*.

Dubousset, J., Asselineau, B., Carlan, L. D. E., Clairand, I., Huet, C., Lacoste, V., & Trompier, F. (2007). A new 2D and 3D imaging approach to musculoskeletal physiology and pathology with low-dose radiation and the standing position: the EOS system. *Radioprotection*, *41*, 227‑252. https://doi.org/10.1051/radiopro

Fang, L., Wang, Z., Chen, Z., Jian, F., Li, S., & He, H. (2020). 3D shape reconstruction of lumbar vertebra from two X-ray images and a CT model. *IEEE/CAA Journal of Automatica Sinica*, *7*(4), 1124‑1133. https://doi.org/10.1109/JAS.2019.1911528

Ferrarini, L., Olofsen, H., Palm, W. M., van Buchem, M. A., Reiber, J. H. C., & Admiraal-Behloul, F. (2007). GAMEs: Growing and adaptive meshes for fully automatic shape modeling and analysis. *Medical Image Analysis*, *11*(3), 302‑314. https://doi.org/10.1016/j.media.2007.03.006

Fleute, M., & Lavallée, S. (1999). Nonrigid 3-D/2-D registration of images using statistical models. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *1679*, 138‑147. https://doi.org/10.1007/10704282_15

Gamage, P., Xie Q., Q., Delmas, P., Xu, P., & Mukherjee, S. (2008). Patient-specific customization of a generic femur model using orthogonal 2D radiographs. Dans *Proceedings - Digital Image Computing: Techniques and Applications, DICTA 2008*. https://doi.org/10.1109/DICTA.2008.94

Gamage, Pavan, Xie, S. Q., Delmas, P., & Xu, W. L. (2010). Diagnostic radiograph based 3D bone reconstruction framework: Application to osteotomy surgical planning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6363 LNCS*(PART 3), 481‑488. https://doi.org/10.1007/978-3-642-15711-0_60

Goswami, B., & Kr., S. (2015). 3D Modeling of X-Ray Images: A Review. *International Journal of Computer Applications*, *132*(7), 40‑46. https://doi.org/10.5120/ijca2015907566

Guenoun, B., Zadegan, F., Aim, F., Hannouche, D., & Nizard, R. (2012). Reliability of a new method for lower-extremity measurements based on stereoradiographic three-dimensional reconstruction. *Orthopaedics and Traumatology: Surgery and Research*, *98*(5), 506‑513. https://doi.org/10.1016/j.otsr.2012.03.014

Haskins, G., Kruger, U., & Yan, P. (2020). Deep learning in medical image registration: a survey. *Machine Vision and Applications*, *31*(1), 1‑18. https://doi.org/10.1007/s00138-020-01060-x

Hatt, C. R., Speidel, M. A., & Raval, A. N. (2015). Robust 5DOF transesophageal echo probe tracking at fluoroscopic frame rates. Dans *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. (S.l.) : (s.n.). https://doi.org/10.1007/978-3-319-24553-9_36

Hosseinian, S., & Arefi, H. (2015). 3D reconstruction from multi-view medical X-ray images - Review and evaluation of existing methods. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(1W5), 319‑326. https://doi.org/10.5194/isprsarchives-XL-1-W5-319-2015

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(4), 679‑688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Jannin, P., Fitzpatrick, J. M., Hawkes, D. J., Pennec, X., Shahidi, R., & Vannier, M. W. (2002). Validation of medical image processing in image-guided therapy. *IEEE Transactions on Medical Imaging*, *21*(12), 1445‑1449. https://doi.org/10.1109/TMI.2002.806568

Jianhua Yao, & Taylor. (2003). Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model. Dans *Proceedings Ninth IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2003.1238644

Kasten, Y., Doktofsky, D., & Kovler, I. (2020). End-To-End Convolutional Neural Network for 3D Reconstruction of Knee Bones from Bi-planar X-Ray Images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12450 LNCS*, 123‑133. https://doi.org/10.1007/978-3-030-61598-7_12

Khameneh, N. B., Vazquez, C., Cresson, T., Lavoie, F., & de Guise, J. (2021). Highly accurate automated patient-specific 3D bone pose and scale estimation using bi-planar pose-invariant patches in a CNN-based 3D/2D registration framework. *Proceedings - International Symposium on Biomedical Imaging*, *2021-April*(Isbi), 681‑684. https://doi.org/10.1109/ISBI48211.2021.9433843

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1‑15.

Knaan, D., & Joskowicz, L. (2003). Effective Intensity-Based 2D/3D Rigid Registration between Fluoroscopic X-Ray and CT. (S.l.) : (s.n.). https://doi.org/10.1007/978-3-540-39899-8_44

Laporte, S., Skalli, W., de Guise, J., Lavaste, F., & Mitton, D. (2003). A biplanar reconstruction method based on 2D and 3D contours: Application to the distal femur. *Computer Methods in Biomechanics and Biomedical Engineering*. https://doi.org/10.1080/1025584031000065956

Liao, R., Miao, S., de Tournemire, P., Grbic, S., Kamen, A., Mansi, T., & Comaniciu, D. (2016). An Artificial Agent for Robust Image Registration. https://doi.org/10.1103/PhysRevX.5.041024

Likar, B., Markelj, P., & Tomaz, D. (2012). A review of 3D / 2D registration methods for image-guided interventions ˇ evic, *16*, 642‑661. https://doi.org/10.1016/j.media.2010.03.005

Lisa Y. W. Tang and Ghassan Hamarneh. (2013). *Medical Image Registration: A Review*. (S.l.) : Medical Imaging: Technology and Applications.

Mahfouz, M., Badawi, A., Fatah, E. E. A., Kuhn, M., & Merkl, B. (2006). Reconstruction of 3D patient-specific bone models from biplanar X-ray images utilizing morphometric measurements. *Proceedings of the 2006 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV'06*, *2*(January), 345‑349.

Markelj, P., Tomaževič, D., Likar, B., & Pernuš, F. (2012). A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis*, *16*(3), 642‑661. https://doi.org/10.1016/j.media.2010.03.005

Markelj, P., Tomaževič, D., Pernuš, F., & Likar, B. (2012). A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis*, *16*(3), 642‑661. https://doi.org/10.1016/j.media.2010.03.005

Melhem, E., Assi, A., El Rachkidi, R., & Ghanem, I. (2016). EOS® biplanar X-ray imaging: concept, developments, benefits, and limitations. *Journal of Children's Orthopaedics*, *10*(1), 1‑14. https://doi.org/10.1007/s11832-016-0713-0

Miao, S., Wang, Z. J., & Liao, R. (2016). A CNN Regression Approach for Real-Time 2D/3D Registration. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2016.2521800

Mitton, D., Landry, C., Véron, S., Skalli, W., Lavaste, F., & de Guise, J. (2000). 3D reconstruction method from biplanar radiography using non-stereocorresponding points and elastic deformable meshes. *Medical and Biological Engineering and Computing*. https://doi.org/10.1007/BF02344767

Mitulescu, A., Semaan, I., de Guise, J., Leborgne, P., & Adamsbaum, C. (2001). Validation of the non-stereo corresponding points stereoradiographic 3D reconstruction technique. *Medical and Biological Engineering and Computing*, *39*(2), 152‑158. https://doi.org/10.1007/BF02344797

Mosca, V. S. (2017). Principles of Lower Extremity Deformity Correction in Children. *Journal of Pediatric Orthopaedics*, *37*(6), S9‑S11. https://doi.org/10.1097/BPO.0000000000001028

Reyneke, C. J. F., Luthi, M., Burdin, V., Douglas, T. S., Vetter, T., & Mutsvangwa, T. E. M. (2019). 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. https://doi.org/10.1109/RBME.2018.2876450

Rodrigues, T. B., Ó Catháin, C., O'Connor, N. E., & Murray, N. (2020). A Quality of Experience assessment of haptic and augmented reality feedback modalities in a gait analysis system. *PLoS ONE*, *15*(3), 1‑19. https://doi.org/10.1371/journal.pone.0230570

Rote, G. (1991). Computing the minimum Hausdorff distance between two point sets on a line under translation. *Information Processing Letters, 38(3),* 123-127.https://doi.org/10.1016/0020-0190(91)90233-8

Roth, S. D. (1982). Ray casting for modeling solids. *Computer Graphic Image Processing*, 18, 109-144. https://doi.org/10.1016/0146-664X(82)90169-1

Salvi, J., Matabosch, C., Fofi, D., & Forest, J. (2007). A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, *25*(5), 578‑596. https://doi.org/10.1016/j.imavis.2006.05.012

Sokooti, H., de Vos, B., Berendsen, F., Lelieveldt, B. P. F., Išgum, I., & Staring, M. (2017). Nonrigid Image Registration Using Multi-scale 3D Convolutional Neural Networks BT - Medical Image Computing and Computer Assisted Intervention − MICCAI 2017, *1*, 232‑239. https://doi.org/10.1007/978-3-319-66182-7

Sotiras, A., Davatzikos, C., & Paragios, N. (2013). Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, *32*(7), 1153‑1190. https://doi.org/10.1109/TMI.2013.2265603

Van De Kraats, E. B., Penney, G. P., Tomaževič, D., Van Walsum, T., & Niessen, W. J. (2005). Standardized evaluation methodology for 2-D-3-D registration. *IEEE Transactions on Medical Imaging*, *24*(9), 1177‑1189. https://doi.org/10.1109/TMI.2005.853240 van Drongelen, S., Kaldowski, H., Fey, B., Tarhan, T., Assi, A., Stief, F., & Meurer, A. (2020). Determination of leg alignment in hip osteoarthritis patients with the eos® system and the effect on external joint moments during gait. *Applied Sciences (Switzerland)*, *10*(21), 1‑14. https://doi.org/10.3390/app10217777

Wybier, M., & Bossard, P. (2013). Musculoskeletal imaging in progress: The EOS imaging system. *Joint Bone Spine*, *80*(3), 238‑243. https://doi.org/10.1016/j.jbspin.2012.09.018

Yao, J., & Taylor, R. (2003). *Assessing Accuracy Factors in Deformable 2D/3D Medical Image Registration Using a Statistical Pelvis Model*.

Youn, K., Park, M. S., & Lee, J. (2017). Iterative approach for 3D reconstruction of the femur from un-calibrated 2D radiographic images. *Medical Engineering and Physics*, *50*, 89‑95. https://doi.org/10.1016/j.medengphy.2017.08.016

Yu, W., Chu, C., Tannast, M., & Zheng, G. (2016). Fully automatic reconstruction of personalized 3D volumes of the proximal femur from 2D X-ray images. *International Journal of Computer Assisted Radiology and Surgery*, *11*(9), 1673‑1685. https://doi.org/10.1007/s11548-016-1400-9

Yu, W., Tannast, M., & Zheng, G. (2017). Non-rigid free-form 2D–3D registration using a B-spline-based statistical deformation model. *Pattern Recognition*, *63*(September 2016), 689‑699. https://doi.org/10.1016/j.patcog.2016.09.036

Yu, W., Zysset, P., & Zheng, G. (2015). Personalized x-ray reconstruction of the proximal femur via a non-rigid 2D-3D registration. Dans S. Ourselin & M. A. Styner (Éds), *Medical Imaging 2015: Image Processing* (Vol. 9413, p. 94133B). https://doi.org/10.1117/12.2082339

Zheng, G. (2011). Personalized x-ray reconstruction of the proximal femur via intensity-based non-rigid 2D-3D registration. *Medical Image Computing and Computer-Assisted Intervention–MICCAI* 2011, 598‑606. https://doi.org/10.1007/978-3-642-23629-7_73

Zheng, J., Miao, S., Jane Wang, Z., & Liao, R. (2018). Pairwise domain adaptation module for CNN-based 2-D/3-D registration. *Journal of Medical Imaging*, *5*(02), 1. https://doi.org/10.1117/1.jmi.5.2.021204

Zollei, L., Grimson, E., Norbash, A., & Wells, W. (2001). 2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. Dans *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. https://doi.org/10.1109/CVPR.2001.991032