AUTOMATED AND COMPUTATIONALLY EFFICIENT JOINT MOTION ANALYSIS USING LOW QUALITY FLUOROSCOPY IMAGES

BY

SOHEIL GHAFURIAN

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Industrial and Systems Engineering

Written under the direction of

Kang Li

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January, 2017

ABSTRACT OF THE DISSERTATION

Automated and Computationally Efficient Joint Motion Analysis Using Low Quality Fluoroscopy Images

by

Soheil Ghafurian

Dissertation Director: Kang Li

The kinematic analysis of joint motion has proven to significantly improve orthopedic surgeries by enhancing surgery assessment, prosthesis design, and pathology diagnosis. This analysis is usually performed by a three to two dimensional (3D/2D) registration of the 3D bone model to a 2D radiographic video from C-arm fluoroscopy imaging machines. However, the practicality of such analysis is undermined due to lengthy and user-dependent 3D/2D image registration algorithms and the high cost of C-arm fluoroscopy imaging machines. Mini C-arm fluoroscopic machines are a more affordable alternative, but the low quality of their images has hindered their use in this application.

In this thesis, we propose a novel 3D/2D image registration algorithm for the kinematic analysis of joint bones using mini C-arm fluoroscopy machines with significantly improved speed, despite their low quality images. This method performs a fast registration through a novel algorithm for quick and high-quality generation of digitally reconstructed radiographs (DRR), which is the bottleneck in such processes. Moreover, the dependency of the results on the user has been reduced as a new feature-based registration algorithm replaced the previously manual initialization phase of the process. This algorithm is able to reach the true registration from within 90 degrees of it, which is a substantial improvement over the existing methods. In

addition, our algorithm performs the registration in significantly reduced time due to a smaller number of generated DRRs.

Acknowledgements

First and foremost, I would like to express my most sincere gratitude to my PhD adviser Professor Kang Li for familiarizing me with this field and helping me push my abilities beyond what I thought I would be able to do before. The freedom and independence he gave me during my PhD course was crucial in my growth as a scientist. I would like to thank Professor Ilker Hacihaliloglu for his positive and intimate support and scientific help, which was utterly constructive for me. I'm also thankful to the members of my committee Dr. Susan Albin and Dr. Myong Jeong, both of whom I have had the honor of being students of.

I would also like to thank the great friends that I have had at Rutgers University especially in the departments of Computer Science and Industrial Engineering. My years as a PhD student would have never been as colorful and memorable without the richness they provided my life with.

Dedication

To my parents, whom I have not seen for six years and whose dream of meeting again has given warmth to my heart everyday since I left them...

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Joint Kinematic Analysis

The goal of most orthopedic surgeries is to restore the natural movement of the joint [51]. Neither achieving nor assessing our performance towards this goal is possible without having a method for precise measurement of joint movement. Such a method is vital to the assessment of orthopedic surgeries, diagnosis of pathological abnormalities, and the design of prosthetic replacements [74, 51, 42, 64, 65]:

- The analysis of joint motion provides a strong tool for evaluating the outcomes of orthopedic surgery by comparing the post operative motion pattern of the patient to a healthy one or the pre-operative one. If such technology is not available, the problems with the surgery will be diagnosed only after they have been intense enough to cause symptoms, which usually requires additional invasive correctional surgeries and are accompanied with pain.

- Kinematic analysis of spine motion could help with the development of quantitative diagnostic tools, which detect pathological abnormalities in motion.

- The ability to measure movement kinematics is helpful in designing prosthetic replacements. By knowing the kinematics of a normal healthy spine, it will be easier to assess the quality of prostheses and improve them toward more natural function.

In the field of mechanics, kinematics, also known as the geometry of motion, is a branch of mechanics which describes the motion of points, bodies (objects), and systems of bodies (groups of objects) without consideration of the causes of motion [71, 6]. In joint kinematic analysis, we are interested in the movement of the bones in the joint. The bones are rigid-body

objects in that the shape of the bone remains the same throughout the movement and needs not to be investigated. Also, the cause of the motion, which is the forced coming from the muscles is out of the scope of the analysis. In other words, we are purely interested in describing the motion regardless of its cause.

The motion of a rigid-body object is nothing but the change of its 3D position with time. Therefore, in order to describe such a motion, we need a function which maps each point in time to a 3D position for the object:

$$f(time) \longrightarrow position$$

The position of a rigid-body object is precisely delineated by six parameters. These parameters are the three translational and the three rotational parameters along and around the 3 axes. In the field of mechanics these three parameters are know as the six degrees of freedom (6DoF). Using the six degrees of freedom, the motion of any rigid-body object could be described perfectly by a function of the form

$$f(time) \longrightarrow (r_x, r_y, r_z, t_x, t_y, t_z)$$

where r_x, r_y, and r_z are the rotation and t_x, t_y, and t_z are the translation parameters around and along the X, Y, and Z axes. The ultimate goal of this thesis is to find this function for every bone in the joint under study.

1.2 Joint Kinematic Analysis Methods

A common method for measuring motion kinematics is the use of skin-mounted markers [74], which is not suitable for orthopedic purposes mentioned above. After such markers are mounted on the skin, the patient is recorded while performing normal motion patterns and the location of the bone is estimated by the location of the markers. The problem with this method is, however, that the motion of the markers are not precise representatives of that of the bone. This is due to the fact that the existence of soft tissue adds a considerable amount of noise to the measurements and violates the rigidity of the object which is one of the assumptions of kinematic analysis of joint movement. Thus, the results from this methods lack a high level of accuracy,

which renders them inappropriate for orthopedic purposes.

Another method which produces exact results, is the use of fiducials implanted on the bone. This method results in the most precise measurements for the kinematics of the bone because the location of the fiducials is completely fixed in relation to the bone therefore representing the rigid-body motion very well. Although this method is commonly used for cadaver studies, its applicability for in-vivo human studies is limited due to its highly intensive nature and the interruption it would cause in the natural pattern of motion due to the pain [51].

To overcome the drawbacks of the above methods, using medical imaging systems is preferred. Since we are interested in the 3D motion of the bones, it would be ideal to have an imaging technology for capturing 3D and dynamic images from the joint movement, however such a technology does not exist. Therefore, to obtain such an image, image registration techniques are used to combine the advantages of different imaging systems.

A common method for kinematic analysis of joint motion using medical imaging is the use of fluoroscopy C-arm machines [51, 8, 1]. In this method, a two-dimensional radiographic video of the joint motion is captured using a C-arm fluoroscopic machine. The video demonstrates the motion of the joint in two dimensions. To calculate the three-dimensional parameters of the movement, a three dimensional model of the bone is registered to each frame of the video based on the assumption that given the same camera and imaging parameters, the 3D bone model must be located similar to the real one to produce a two dimensional projection which matches the radiographic frame captured in the clinical setting. After this registration is done perfectly for every frame in the video, the 3D position of the bone is at hand for the movement (Figure 1.1). The three dimensional model of the bone is usually obtained from a segmentation of the bone in computed tomography (CT) image series, which provide the 3D information of the bone, but are not able to capture dynamic videos of the motion.

While the above method is commonly used, it is still a challenging problem because the precision of the method depends on the quality of the image registration. The three-dimensional to two-dimensional (3D/2D) image registration used in this method is by nature a hard problem due to the loss of information in the projection of a 3D object to the 2D space. This is exacerbated by the low quality of fluoroscopy images the multi-modality of our problem. In the following section, we introduce the problem of image registration and then we will review

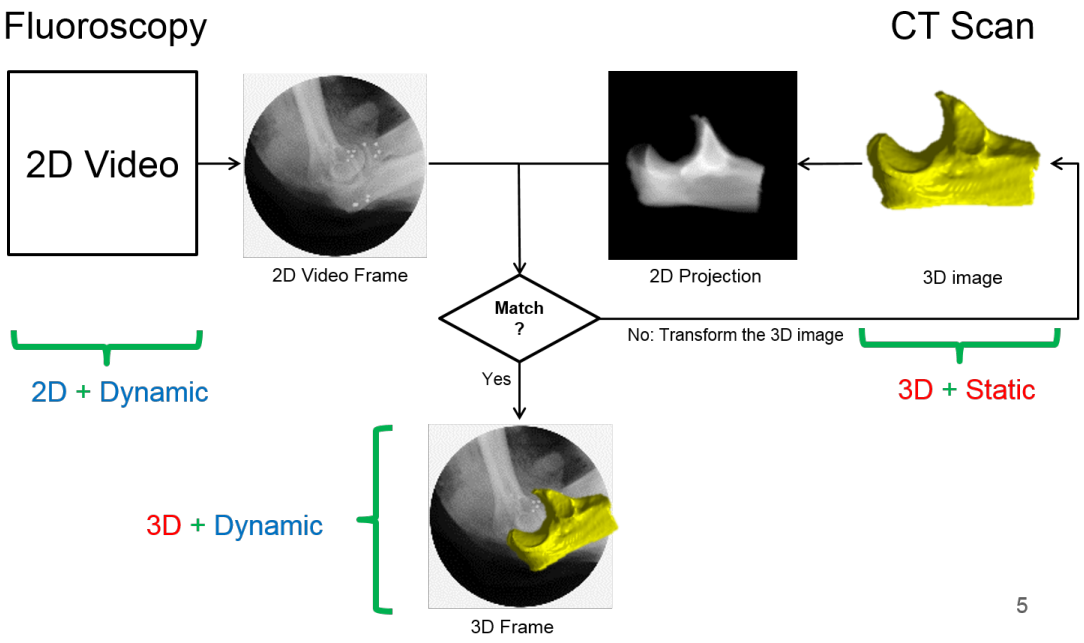


Figure 1.1: To obtain a three dimensional (3D) and dynamic model of bone movement, image registration used to combine the 3D nature of the CT imaging with movement information obtained from the fluoroscopy machine in clinical settings. A 3D volume of the bone is segmented from CT image series and is registered to every frame of the fluoroscopy video to provide the 3D position of the bone at each time during the motion. This registration is done by finding the 3D location of the bone which results in the most similar 2D projection, namely digitally reconstructed radiograph, to the fluoroscopy frame. Finding this location is an iterative optimization process over the six-dimensional space of possible locations.

some of the methods for joint kinematic analysis using the method mentioned above.

1.3 Image Registration

Image registration, also known as image matching, is the process of transforming different images into one coordinate system [12]. The two images might be obtained from different sensors, times, depths or view points. This unification of the coordinate systems is performed by finding a geometric transformation, which transforms the coordinate system of one of the images (moving, target, sensed, or the subject image) to the coordinate system of the other image (the fixed, source, or reference image). With such a transformation at hand, the coordinate system of any point in either of the images could be calculated on the other. Depending on the nature

of the geometric transformation, the type of the images, and the registration process, image registration falls into different categories.

1.3.1 Types of Image Registration

Image registration methods could be different in terms of the nature of the geometric transformation which converts the coordinate systems into each other. In deformable registration, the transformation between two images could be non-linear which results in a distortion and reshaping of the moving image. However, in affine image registration, straight lines and planes are preserved and parallel lines and planes will be still parallel after the transformation [7]. If the scaling factor of an affine registration is eliminated, the registration will be a rigid-body one. In a rigid body image registration the transformation of the image will be confined to translation and rotation.

Registration methods could differ based on the modality of the images. If both the reference and the target image are captured using the same kind of image sensor, the registration is called single modality image registration. However, if the two images are obtained from two different kinds of image sensors, the registration is a multi-modal one. Multi-modal image registration methods are usually more complicated than the single-modality ones and require more sophisticated solution processes.

Image registration methods could be either intensity-based or feature-based. Intensity-based image registration is concerned with aligning the target image to the reference image by comparing the intensity patterns in the two images [27]. Feature-based image registration methods, on the other hand, accomplish this task by comparing features extracted from the two images. Such image features include landmarks, contours, lines, or statistical features [27].

The registration of images could be performed with or without user interaction. Manual registration frameworks provide the user with the tools to perform the registration visually. Interactive registration methods rely on the user for certain steps of the registration. Semi-automatic registration performs the main part of the registration automatically, however it relies on the user for the initialization of the process or verification of the results, while an automated registration framework is completely independent of the user [27]. A fully automated registration of images is generally ideal, but not always possible.

Image Registration Based on Image Dimensionality

The most significant categorization of image registration methods stems from the dimension of the images. In two-dimensional to two-dimensional (2D/2D) image registration, both the target and the reference images are two dimensional (2D) images. In three-dimensional to three-dimensional (3D/3D) image registration, both of these images are three-dimensional, which is also called a volume. In a three dimensional image, the value of each pixel represents the sampled intensity value of a 3D scene rather than a 2D sensor. Such three dimensional pixels are also called voxels. In the context of rigid-body image registration, a 2D/2D image registration involves two in-plane translations and an in-plane rotation and a 3D/3D image registration entails three translational and three rotational transformations. The registration of a 3D image to a 2D one is another class of important image registration methods, which is slightly different from the two aforementioned ones (Figure 1.2).

In three-dimensional to two-dimensional (3D/2D) image registration, it is the 2D projection of the 3D target image which is aligned with the 2D reference image and not the 3D image itself. The aim of 3D/2D image registration is to find the location of the volume in space at the time when the reference 2D image was captured. The result of the registration will be a rigid geometric transformation, which transforms the coordinate system of the 3D volume to that of the real object at the time when the 2D image was captured. In this approach the camera specifications are measured and a virtual simulation of the imaging gantry is set up. The premise of this method is that if the projection of the 3D image is successfully aligned with the 2D reference image, the location of the 3D image in space must coincide with that of the real object at the time of imaging. Therefore, this kind of registration involves searching the space of possible geometrical transformation for the most similar 2D projection to the 2D reference image.

1.3.2 Image Registration as an Optimization Problem

Image Registration could be considered as an optimization problem. As explained in section 1.3.1, 3D/2D image registration aligns the coordinate system of the reference image and the 2D projection of the 3D target image by maximizing the similarity of the two 2D images.

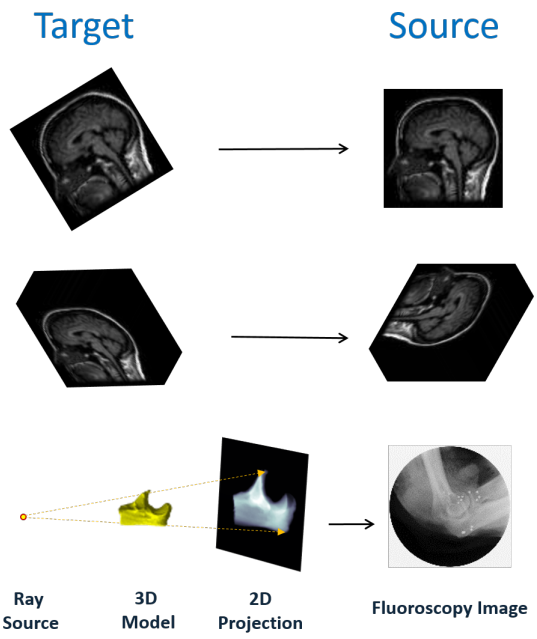


Figure 1.2: The goal of image registration is to find the transformation parameters that match the coordinate system of the target image with that of the source image. In two-dimensional to two-dimensional (2D/2D) image registration, this transformation consists of three parameters, two in-plane translations and on in-plane rotation (top). In three-dimensional to three-dimensional (3D/3D) image registration, all the six degrees of freedoms (three translational and three rotational parameters) need to be found (middle). In three-dimensional to two-dimensional image registration, like the 3D/3D case, it is the six degrees of freedom that are sought after. However it is a 2D projection of the 3D target image that needs to match the source image and not the target image itself (bottom).

Therefore, any registration problem can be looked at as a maximization problem, in which the objective function to be maximized is the similarity of the two images.

As a non-linear optimization problem, 3D/2D image registration does not have a closed form solution. Therefore, this problem needs to be tackled with iterative optimization protocols. Such optimization methods fall into two general categories of derivative-based and derivative-free methods [55]. Since the similarity of the DRR and the fluoroscopy image is not a straightforward mathematical function of the position (six degrees of freedom) of the volume, the derivative of which is easily calculable, it needs to be tackled through derivative-free methods, which estimate the direction of the gradient through an assessment of current solutions.

In any iterative optimization method, four elements need to be determined:

- objective function,
- search strategy,
- initialization,
- stopping criteria.

The objective function is a function which takes in a solution and returns the fitness of the solution. The goal of optimization is to find the optimum objective function. Search strategy is the process deciding which neighborhood of the current solutions to search for optimum solutions. Initialization is the protocol according to which we decide which solutions to start the optimization algorithm with. And the stopping criteria determines when the search has ended and the current best solution should be returned as the optimum answer.

As an optimization process, any 3D/2D image registration needs to specify the equivalent of the aforementioned optimization elements. The objective function of such a method includes calculating the similarity of the source image and the projected 3D image. Also, the mechanism for the projection of the 3D image onto the 2D image plane is an important part of the implementation of the objective function. Search strategy is usually taken from the existing derivative-free optimization methods. Initialization of the process is usually done visually using applications that provide the tools for this job. Therefore, any image registration process needs to determine the details of the following steps:

- 2D image similarity measure,
- volume projection mechanism,
- search strategy,
- and initialization.

1.4 Related Work

For the in vivo kinematic analysis of the joint motion, to avoid the inaccuracies associated with the skin-mounted markers and the invasiveness of the bone-implanted fiducials, 3D/2D image

registration has been proposed by Pickering et al. [21, 18, 45, 38]. In this method, a video of the normal motion of the joint is captured through fluoroscopy imaging, which is a perspective X-ray imaging system. Thereafter, a 3D model of each joint bone is registered to each fluoroscopy frame in the video. The 3D bone model is reconstructed through a segmentation of the bone of interest in a Computed Tomography (CT) image series. The result of this process is a 6-degree-of-freedom transformation delineating the location of the bone for each frame of the video, which shows the location of the bone when that frame was captured. Fregly et al. [21] have performed a theoretical analysis of the accuracy of this approach [51].

One of the main challenges of using 3D/2D image registration for in vivo bone kinematic analysis is the inaccuracy of results. Specifically, finding the parameter for out-of-plane translation is hard [21]. Out-of-plane translation is the movement of the volume in the direction perpendicular to the image plane. Since this movement manifests only as a minimal size change in the 2D projection, it is difficult to retrace the exact amount of this movement.

You et al. [74], studied the in vivo kinematics of the knee joint by the use of a bi-planar fluoroscopy machine. A bi-planar fluoroscopy machine consists of two ray sources and two sensors (two C-arms) the focal axis of each pair of which is perpendicular to the other (Figure 1.5). These two arms, capture the motion of the subject independently, the results of which is two fluoroscopic videos of the movement from two different points of view. The advantage of using a bi-planar system is that the out-of-plane translation of each video is represented as an in-plane translation in the other, resulting in much more precise results compared to a single-plane fluoroscopy [74].

In their study, You et al. [74] attained a precision of 2.5°for rotation and 0.8 mm for translation for the in-vivo kinematics of the knee. For the optimization method of the registration problem, they used the Downhill Simplex method, which maximized the similarity of the DRR and the fluoroscopy image by increasing the normalized cross correlation of the pixel intensities in the two images. The Optimization method was iteratively restarted from the optimized solution to escape a possible local optimum. Prior to the calculation of the similarity measure, both the DRR and the fluoroscopy image are edge enhanced. This enhancement is done by calculating the edges of the images (Sobel edge detection) and adding it the original image. Also, a Quad-tree-based method of calculating the cross correlation was employed to avoid the

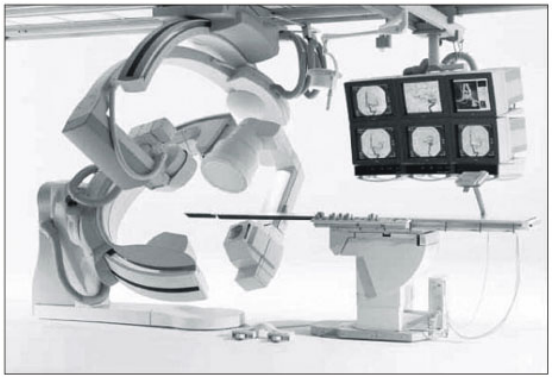


Figure 1.3: A bi-planar fluoroscopy C-arm machine. Such imaging machines have two image sensor and ray sources which capture two images of the subject with perpendicular focal axes. The use of bi-planar C-arm machines is advantageous over the single-plane alternatives due to the fact the the out-of-plane translation of each image series is in-plane for the other. However, such radiography machines are expensive and the extra imaging arm confines the subjects natural range of motion.

high number of iterations which is associated with the cross correlation. The similarity measure was optimized by the downhill simplex optimization method. The DRRs were estimated by a fast rigid transformation of the volume using the texture memory of the computer graphical processing units (GPU) and sampling of planes in the volume so that they are parallel to the image plane image plane. A summation of the sampled planes then results in an estimation of the DRR, however the perspective nature of the DRR is ignored.

Although the use of bi-planar fluoroscopy improves the results of the analysis, it is not always practical. This is due to the high cost of such machines and also the fact that the existence of an extra C-arm confines the movement of the subject substantially, which might make a natural motion of the subject impossible. Therefore, it is highly desirable to have methods which are able to extract good enough results from single-plane fluoroscopy C-arm machines.

Pickering et al. [51] targeted the problem of out-of-plane translation by using a new image similarity measure called Cross-Cumulative Residual Entropy similarity measure. This similarity measure is introduced by Wang et al. [68] and is used in different image registration frameworks [52]. Pickering et al. used this method on the knee joint and obtained improved accuracy, faster convergence, and larger initial difference in the results compared to the widely

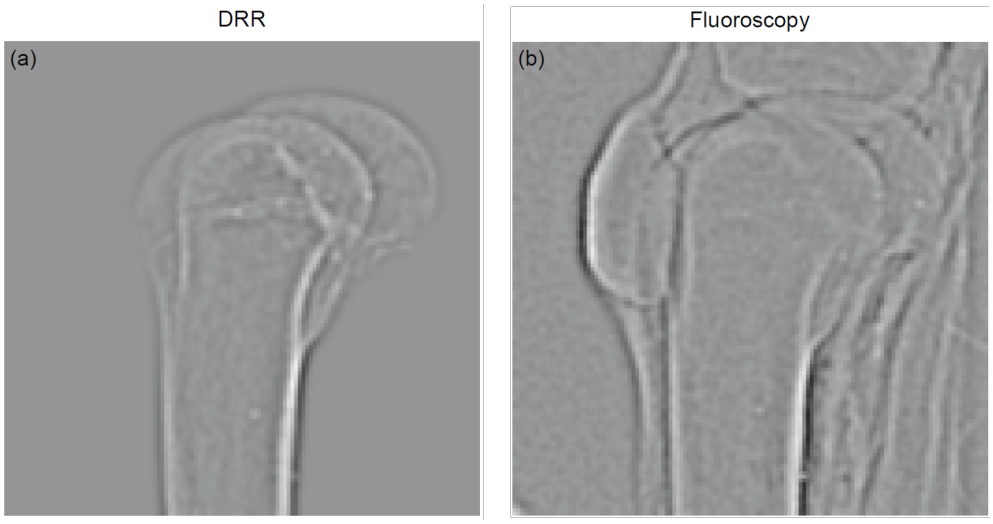


Figure 1.4: Pickering et al. [51] use an edge-enhanced version of the DRR (a) nad the fluoroscopy image (b) for the calculation of the similarity measure. The edges are calculated by applying a Log of Gaussian (LoG) to both images. The variance of the LoG follows a coarse-to-fine pattern, because sharp edges in the beginning of the algorithm increase the probability of getting stuck in local optima.

used mutual information similarity measure. This similarity measure was applied to an edge-enhanced version of the images, which were produced through the Log of Gaussian (LoG) filter (Figure 1.4). The use of such edge enhancement adds to the accuracy of the final results, but causes a higher possibility of the algorithm getting stuck in local maxima. To avoid this problem, a coarse-to-fine approach was adopted which narrowed the LoG filter kernel as the solution process advanced.

To avoid the time intensity of DRR generation, in this method, the DRRs were estimated by summing the volume voxel intensities in the Z direction after the geometric transformation was applied to the volume. Moreover, the resolution of both the CT and the fluoroscopy images was reduced to avoid lengthy calculation times. This is not an ideal strategy for DRR generation because the fluoroscopy C-arm is a cone beam camera model, which means the projection is perspective as opposed to parallel since the ray source is a point and the X-rays converge from it. Therefore, such a method for DRR generation deteriorates the results of the algorithm potentially. Also, the registration error in the results was partly attributed to the reduction in resolution due to lack of enough processing power.

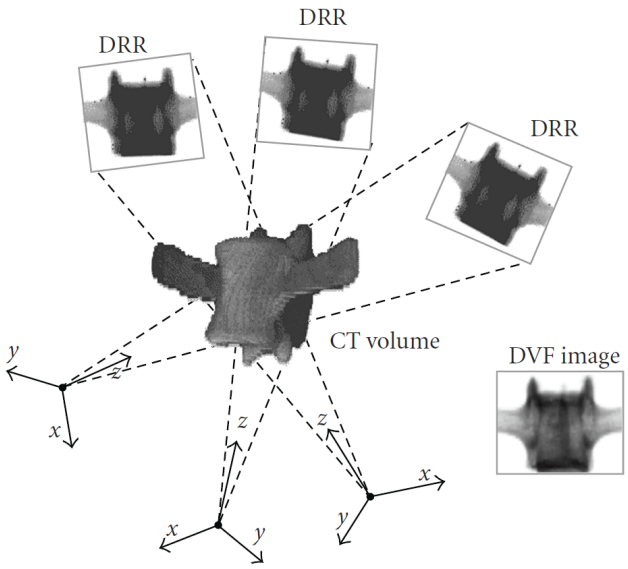


Figure 1.5: A rotation of the ray source instead of the 3D volume results in a significant reduction in DRR generation time [8]. This is an approach which is commonly used for DRR generation.

Bifulco et al. [8] applied a similar method for the kinematic analysis of sheep vertebra. The results showed an error of 0.1°for rotation and 1mm for in-plane translation. For out-of-plane translation, however, the error was 10mm. In this method, the cross correlation is used as the image similarity measure. The search strategy is based on a likelihood map, which at each iteration samples the neighborhood of interest in the solution space and confines the neighborhood of interest to the vicinity of the solution with highest similarity measure. This neighborhood is sampled with a higher resolution in the next iteration until the algorithm is converged towards an optimum solution. The problem with this search strategy is that depending on where the sampled solutions are located, there is a high risk of getting stuck in local optima. To shorten the generation of DRRs, the resolution of the DRRs is reduced in the beginning of the process and is gradually increased as the solution process continues (Figures 1.5 and 1.6). This exacerbates the chance of getting stuck in local optima as the initial similarity values are not exact.

Tersi et al. [64] verified the applicability of such methods for the kinematic analysis of elbow in-silico. This method was in-silico in that instead of real clinical fluoroscopic X-ray images, two dimensional projections of the 3D model of the elbow joint were used. In this

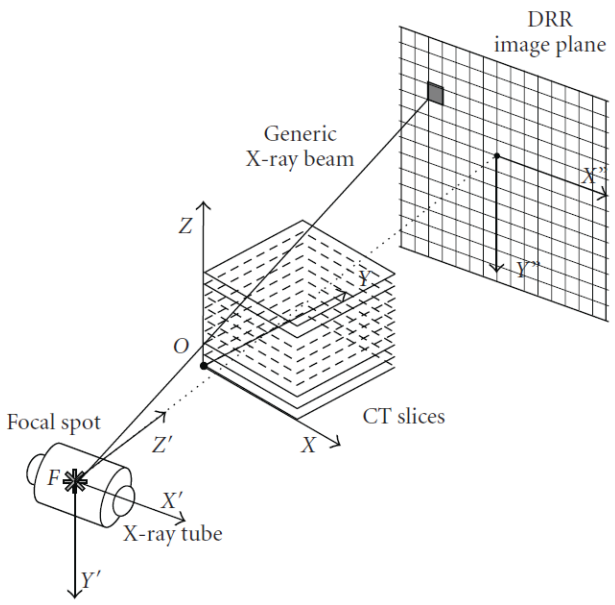


Figure 1.6: The Digitally Reconstructed Radiograph (DRR) is a simulation of the fluoroscopy imaging process (Picture from [8]). To generate DRRs, the image plane and the x-ray source are set-up after the real x-ray C-arm and the value of each pixel on the DRR is determined by the average intensity of the voxels that the ray to that pixel passes through.

study, instead of a 3D volume, a 3D surface model of the elbow joints is used, which only carries the information of the surface of the bone and not the inside intensity values. Consequently, instead of DRRs, which integrate and project the intensities of the voxels inside the volume, a simple projection of the surface models was employed, which shows on the 2D contours of the projected surface model. Such a projection is fast to generate, but lacks the inside information of the 3D model, which reduces the precision of the results in real clinical applications.

Weighted edge matching score (WEMS) is an image similarity measure, which is introduced by Tsai et al. [65]. In this method, the edges of the fluoroscopy image are extracted using the canny edge detection method and edges which are shorter than a certain threshold or omitted. The remaining edges are then iteratively dilated and added to the canny edge Image. To calculate the WEMS for each DRR, the edges of that DRR are extracted using canny edge detection and weighted according to the lengths of the edges. Thereafter, the two preprocessed images (DRR and fluoroscopy) are multiplied pixel-wise and the sum of the pixels values of the result over the sum of the edge weights of the DRR is considered as the WEMS (Figure 1.8).

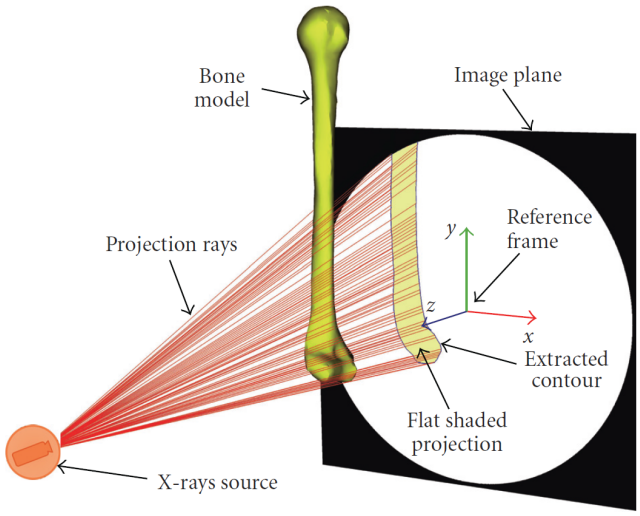


Figure 1.7: Tersi et al. [64] used a 2D projection of the bone surface rather than the DRR. The process of generating such an image is significantly fast than a DRR, but lacks the information from inside the bone. Due to the low quality of the fluoroscopic images, such information is essential for a precise registration of the 3D model [74].

The optimization technique which is used for the maximization of the similarity in this study is genetic algorithm. Genetic algorithm is an iterative population-based optimization method, which does not depend on the gradient of the objective function. Although this method is capable of reaching the solution in this case, due to being population-based, it requires the assessment of many solutions and consequently the generation of many DRRs, which make the process time-intensive. The solution time or the DRR generation method in this study has not been reported, but the precision of the results is reported to be 0.77 millimeters for in-plane translation, 3.06 millimeters for out-of-plane translation, and 1.13°for rotations.

All of the methods mentioned above and other methods in these category face four major challenges that have kept these methods from being more commonplace in the field. These challenges as follows.

1. Cost:

   A full-size fluoroscopy machine is highly expensive and could cost hundreds of thousands of dollars. This high costs makes it impossible for the majority of labs to benefit from using such methods. Practically, only few labs have the possibility of implementing one of these methods.

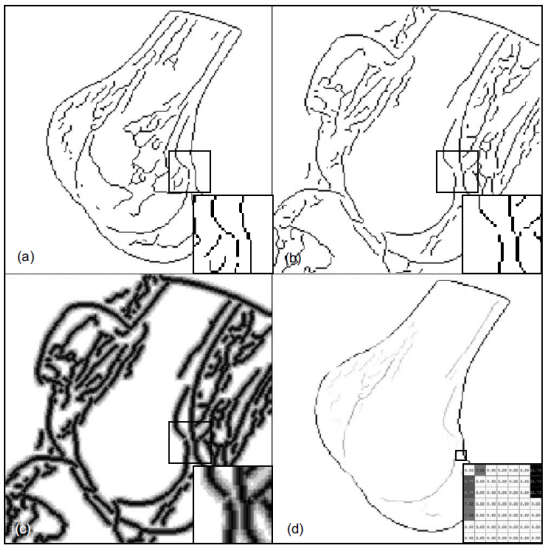


Figure 1.8: Image processing steps for WEMS [65]. The edges of the DRR (a) and the fluoroscopic image (b) are detected through a canny edge detection algorithm. (c) The edges from the fluoroscopy image are thresholded dilated so that the edge information are spread in the vicinity of the lines. (d) The edges from the DRR image are weighted according to the 8-neighborhood length of the edges.

2. Speed:

   3D/2D image registration, which is part of all of the methods mentioned above, is generally a computationally intensive problem. As mentioned before, this class of registration methods is an optimization problem which searches the space of solutions for the maximum similarity between the 2D image and the projection of the 3D volume. Since the gradient of the objective function in this problem is not available, this problem is usually tackled using iterative heuristic optimization methods, such as genetic algorithm, which require the assessment of the objective function for many solutions. The assessment of each solution requires the project of the 3D volume onto the DRR image plane, which is a complex procedure given the cone-beam structure of imaging in a C-arm. Therefore, the high number of generated DRRs results in a lengthy solution process.

3. Sensitivity to manual initialization:

   Image registration algorithms used in the above methods are able to converge to the correct solution only from a limited distance from the answer. The reason is that due

to the complexity of the solution space, if the registration starts from an initial solution which is far from the answer, the algorithm might get stock in local optima along the way and never reach the globally optimum solution. This necessitates an initial rough registration of the volume to a relatively small distance from the answer, which is done by the user. This manual initialization is not always easy as the 2D projection of the volume could be misleading for human eyes. Also, even if the convergence range of the algorithm is known, one cannot know whether they are close enough to the answer for the algorithm to converge or not, because the answer is not known. Therefore, one has to initialize the algorithm, run it, and wait for the final result to decide whether the initial solution was close enough. This results in multiple runs of the algorithm for each subject, which not only adds to the processing time, but also requires the feed back of the user during the process.

4. Precision:

   Getting a precise location for the bone is the ultimate goal of joint kinematic analysis. This is challenging by nature because through a projection of a 3D volume to an image, one dimension of the data is eliminated and it is hard to reconstruct the 3D information by definition. Moreover, the quality of the 2D image is not always good enough to give us all the detail which might help improve the results. Even in good quality radiographs, the overlapping of the bones in the joint or the effect of soft tissue could cover critical information, leading to poor results in terms of precision.

1.5 Contribution

In this study, we propose a method for the kinematic analysis of joint bones. This method is an improvement over similar existing methods in several ways:

- The use of a mini C-arm fluoroscopy machine over a regular full-size one,
- A substantial decrease in the registration time through a novel method for a fast generation of the digitally reconstructed radiographs,
- A fast and novel method for the initialization of the registration process, which minimizes

the dependence of the results on the user,

- An improvement of the final registration results by the use of a new image similarity measure.

The existence of a method for kinematic analysis using the mini C-arm fluoroscopy machine is highly desirable, but challenging. The mini C-arm fluoroscopy machine is a smaller version of the regular fluoroscopy C-arm used in the methods discussed in section 1.4. The mini C-arm is much more widely available in clinical settings due its significantly lower cost and its mobility compared to a full size fluoroscopy machine. Also, the amount of radiation is less than a full size machine which reduces the risk of complications associated with radiographic imaging. These factors make it more enticing to use the mini C-arm for kinematic analysis. However, all existing methods, to the best knowledge of the author, have used the full size fluoroscopy machine due to the problems associated with the mini C-arm. The images obtained from the mini C-arm are considerably noisier than a full-size fluoroscopy machine. In addition to the high amount of noise in each image, a temporal noise is also associated with each frame. This makes the already hard task of image registration highly challenging for mini C-arms. Also, in mini C-arm fluoroscopy machines, the intensifier has a smaller diameter, which causes a larger part of the bones to be left out, leading to a smaller amount of data to perform the registration with. In this thesis, we have introduced a method for kinematic analysis of the neck spine using images obtained from a mini C-arm.

In the next chapter (chapter 2), we introduce a novel method for a fast generation of digitally reconstructed radiographs (DRR). DRRs are the 2D projections of the 3D bone model, which simulate the clinical fluoroscopy images. The generation of the DRRs is highly computational-intensive, which causes long registration times as the registration process involves the generation of many DRRs. This makes kinematic analysis through image registration less practical due to time constraints in clinical settings. Our method is based on the idea of preprocessing all possible pixel intensities in the DRR image prior to DRR generation. Using this method, we were able to reduce DRR generation time down to two orders of magnitude in our experiments. The amount of processing and the memory requirements showed to be practically small and the method supports any possible pose of the 3D bone model, which is an improvement over all of

the existing methods for quick DRR generation.

In chapter 3, we introduce a new method for general 3D/2D image registration. This method uses the probability density function of the image gradient directions as a feature for image registration and is able to perform the registration with significantly less number of required DRRs, which helps shorten the registration time significantly. A main advantage of this method is that it is robust to the initial position of the bone, a property which is an important improvement over existing methods. This method was able to converge to the answer with an initial solution within $\pm 90^{\circ}$ of the true registration. This property makes this method ideal for eliminating or minimizing user interaction for the registration process, if the bones of interested are isolated from the rest of the fluoroscopy frame image.

Finally, in chapter 4, we suggest a complete process flow for the 3D/2D registration of the bone volume to the fluoroscopy video frames. In this chapter, the method introduced in chapter 3 is used for an initial registration of the volume and also enhancing the inter-slice registration. The isolation and reconstruction of the bone of interest, which is essential to the usability of this method, is performed by a deformable registration of an atlas fluoroscopy image to each subject frame. The atlas is a fluoroscopy frame, for which the registration and therefore the registered DRRs is known. The registration results from this method are then passed on to an intensity based registration method which performs image matching using edge information. This method is then able to provide fine registration of the volume, which provides the necessary information for the kinematic analysis of the bone and therefore joint movement.

Chapter 2

Fast Generation of Digitally Reconstructed Radiographs through an Efficient Preprocessing of Ray Attenuation Values

2.1 Introduction

Digitally reconstructed radiographs (DRR) are synthetic radiographic images, which play an important role in medical three-dimensional to two-dimensional (3D/2D) image registration. The generation of DRRs, due to its complexity, is the bottleneck in 3D/2D image registration algorithm making them time-consuming to the point of impracticality. Such registration algorithms are usually optimization problems which aim at maximizing the similarity of the clinical radiographic 2D image and the perspective projection of the 3D volume with regard to the 3D position of the volume in the space. These problems are usually attach through iterative optimization algorithms, which require the generation and assessment of many solutions, namely 3D positions of the volume. The assessment of each solution requires the generation of one DRR, which reveals the criticality of a fast DRR generation mechanism to perform a quick 3D/2D image registration.

In this chapter, we propose a novel approach for the generation of DRRs which is able to reduce the complexity of DRR generation process. This method is an improvement over similar existing approaches in that:

- it is able to support any 3D position of the volume;
- it makes more efficient use of the memory
- require no additional calculations for the initialization of the algorithm.

We test the effectiveness of our method using three human bones at different quality levels and compare the results to the traditional method for DRR generation, namely ray casting method.

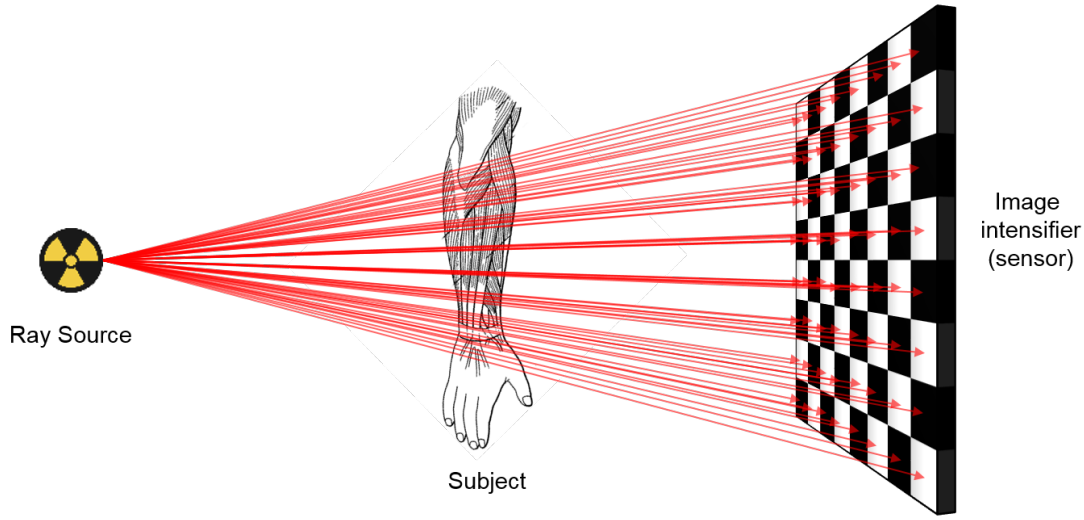


Figure 2.1: DRRs are simulations of perspective radiographic images, where the intensity value of each pixel in the radiographic image is determined by the attenuation value of the rays from the ray source to the corresponding sensor. The rays are attenuated based on the density of the subject tissues they pass through.

2.1.1 The Digitally Reconstructed Radiograph

The DRR is a simulated radiographic image and is essential in several medical applications, where 3D/2D image registration is involved. Examples of such applications include preoperative surgery planing, intra-operative therapy guidance [58], patient placement for radiotherapy planning and treatment verification [9, 25], radio-surgery [48], cranial neurosurgery [41], neuro-interventions [35, 37], spinal surgery [40, 69], orthopedic surgery [34], and aortic stenting procedures [69, 11, 50, 58].

The DRR is obtained through a simulation of the real world radiographic imaging process. A real world radiographic imaging machine consists of a ray source and a sensor, which is also known as the image intensifier (Figure 2.1). Rays, which are emanated from the ray-source, pass through the subject and, depending on the intensity of the tissues they pass through, get attenuated by different amounts. The image intensifier captures the attenuated intensity of these rays and translates them to 2D image intensity data. The intensity of each point on a radiographic image is therefore proportional to the attenuation value of the ray radiating to the corresponding point on the image intensifier. The DRR is calculated by simulating this imaging process and has the same size and spacing of the real radiographic image.

Fast calculation of DRRs is desirable, because a multitude of them need to be produced in medical applications. The problem of 3D/2D image registration is usually tackled by iterative or population-based optimization algorithms, which aim at maximizing the similarity of the DRR to the clinical radiographic image by searching in the space of possible rigid 3D transformations. In such algorithms, the objective function is the similarity measure of the two aforementioned images and each solution is a transformation in space. To evaluate the objective function for each solution, one DRR of the accordingly transformed volume should be calculated. Therefore, to solve such problems, hundreds of DRRs need to be produced, which due to the complexity of DRR generation using conventional methods, make the solution process time-consuming to the point of impracticality.

2.1.2 DRR Generation Ray-Casting and its Complexity

The traditional method for the generation of the DRR is ray-casting. In this method, the intensity of each pixel in the DRR is calculated by forming a ray from the ray-source to that pixel. The location of the ray-source with regard to the DRR is found from a prior camera calibration phase. The details of the camera calibration method can be found in chapter 4. To simulate the attenuation of the ray in the real world, the attenuation value of each ray is found by integrating the intensity of volume voxels along the ray according to Beer's Law [33],

$$I(ij) = R_0 \exp \int_l \mu(v) dv,$$

where $I(ij)$ is the intensity of pixel ij; R_0 is the original intensity of the ray; l is the length of the ray; and v is the intensity of the voxels throughout the ray. This integration is approximated through a sampling and averaging of point intensities along the ray. To reverse the quantization of the volume voxels, the value of each point on the ray is obtained by performing tri-linear interpolation on the eight closest voxels. To save time, only the portion of the ray which passes through the volume (the effective length of the ray) is considered for sampling (Figure 2.2).

The ray casting method is a complex one, which makes the use of 3D/2D image registration less attractive due to practical time constraints. The complexity of this method is $O(ds^2)$, where s is the size of the DRR and d is the number of sampled points on each ray. To sample each

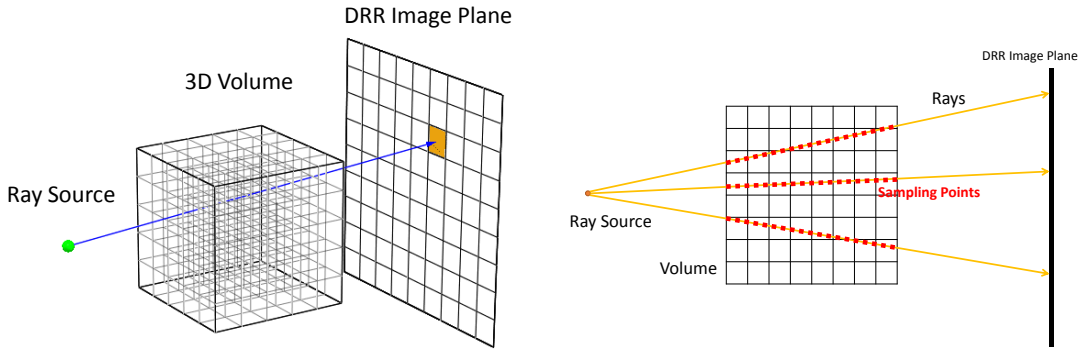


Figure 2.2: The traditional method for DRR generation is ray casting, which is computationally intensive procedure. For generation of DRRs, the image plane and the ray source are set up after the characteristics of the imaging machinery. Left: in the traditional ray casting method, the value of each pixel in the DRR is calculated as the attenuation value of the ray from the ray source to that pixel. Right: The attenuation value of each ray is calculated by sampling the intensity of the volume along the ray at equal distances or proportions. Each intensity sampling requires accessing eight voxel intensities.

point on the ray, the value of eight closest voxels needs to be read, each of which is a memory access operation, which is time-intensive by nature.

2.1.3 Alternative DRR Generation Methods

To produce DRRs in shorter time, alternative methods to ray-casting have been suggested. This speed, however, comes at the price of image quality or memory. This is due to the fact that this class of algorithms aim at reducing the DRR generation time by either a simplification of DRR generation process or preprocessing. Although the loss of quality might not be important in some applications, it is harmful to 3D/2D image registration algorithms, which rely on the detailed info of the 2D images to tackle an already challenging problem. Some of such algorithms are as follows.

Splatting is a projection method, adopted from computer graphics, which reduces the DRR generation time by projecting voxel intensities onto the DRR image plane rather than calculating the rays attenuation value [70]. Through the use of this method, the complexity of DRR generation depends solely on the number of volume voxels rather than the number of point samples on the rays and the size of the DRR. Although this method results in significant reduction in time, the quality of the DRR suffers. Splatting causes an aliasing effect on the DRR which is

the result of the projective magnification of the distance between the volume voxels, for which no intensity is sampled. This aliasing effect is substantially worsened when the volume is closer to the ray source. This distortion is usually addressed by smoothing, which results in a blurry image and deterioration of image quality.

The filtering of the DRR in splatting is not only time-consuming, which to some extent defeats the purpose of this algorithm, it still does not recover the DRR fully. Therefore, Bilkfellner et al. [10] suggest anti-aliasing the DRR through a stochastic distortion of the voxel or ray source location. The purpose of this method is to alleviate the aliasing effect by adding a small amount of randomness to the process. In other words, to randomly smooth the DRR at the time of generation rather than assigning time to it afterwards. Although this method is quicker than a filtering of the DRR, it still results in considerably lower quality images than ray casting. In my replications of the method, when the volume was close to the ray source, the resulting DRR from splatting methods was never of enough quality for a successful registration.

Russakoff et al. [58] suggest a data structure, namely attenuation fields (AF), for storing the pre-calculated values of ray integrals prior to DRR generation. In this method, each ray is uniquely identified by its intersection points with two parallel planes on two sides of the volume. Any combination of two points on the two planes forms a ray. Therefore, the points on the two planes are sampled and their corresponding attenuation values are pre-calculated. When generating a DRR, the attenuation value of any new ray is calculated by performing quadri-linear interpolation on the 16 closest pre-calculated rays, which is significantly less computationally intensive than sampling volume intensities along the ray. This method is able to produce good quality DRRs in short time, however, each attenuation field look-up table supports a limited number of rays resulting in a hard-to-calculate limitation on the range of transformations for the volume. Using additional look-up tables tove cover the out-of-range motion is inefficient and impractical due to the large amount of memory required for each look-up table. Also, the look-up tables share a large number of rays, which leads lengthy and unnecessary precalculation time. Moreover, the AF requires relatively complicated geometric calculations to determine the size and the location of the two planes.

The movement limitations of the AF were improved in progressive attenuation fields (PAF) suggested by Rohlfing et al. [56]. In this method, rather than two planes, each ray is identified

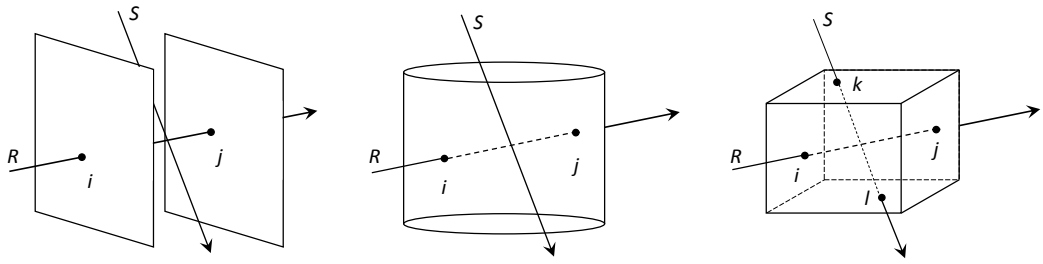


Figure 2.3: The use of two planes in Attenuation Field [58], or a cylinder as in the Progressive Attenuation Field method [56], for the identification of rays going through the volume results in limitations on the transformation of the volume. This is due to the fact that these two structures do not support all of the rays that pass through the volume. (Left) While the ray R has two intersections with the two planes and is captured by the attenuation field method, the ray S, which is closer to parallel with the planes is not supported. (Middle) The ray R has two intersections with the surface of the cylinder, whereas the ray S, which is close to parallel with the axis of the cylinder is not captured. (Right) The use of a cube, as suggested in this paper, assures that every ray passing through the volume is captured and precalculated, which results in a total coverage of possible transformations of the volume.

by its intersection points with the surface of a cylinder encompassing the volume. The use of a cylinder results in the coverage of a full 360°rotation of the volume around the axis of the cylinder. However, rotation around the other axes and also translation along the axis of the cylinder are still limited. Although, it is reasoned that such movements are rare in clinical circumstances, this limitation is problematic when registration is performed in a geometrical coordinate system rather than a clinical one [23], such as our registration framework. Also, not every pair of points on the surface of a cylinder correspond to a ray which passes through the volume, which results in inefficient use of memory because of assigning memory to rays that are not attenuated at all.

The use of a sphere instead of a cylinder could potentially remove the constraint of transformations, but it is not straightforward to implement. If the points on the sphere surface are sampled according to a spherical grid 2.3, the rays resulting from the points closer to the poles will be much more akin to each other than the ones far from them. This will result in an inconsistent sampling of the rays and therefore a waste of memory. Hence, a sampling grid is required where the distances between the sampling points are equal. A geodesic grid on the surface of the sphere (Figure 2.4) satisfies this requirement. Using a geodesic grid instead of

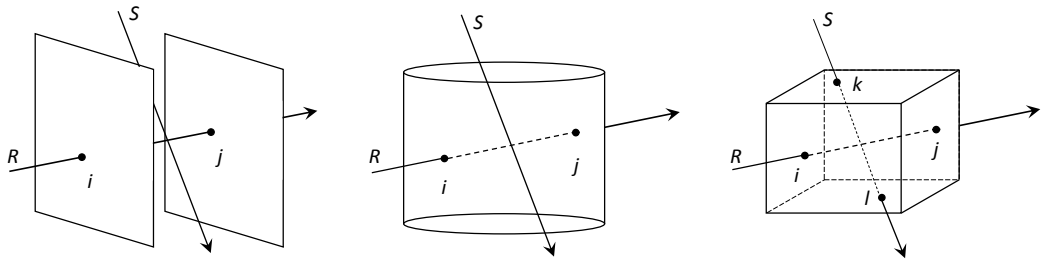


Figure 2.4: Russakof et al. [58] introduced the attenuation fields, in which each ray (R) going through the volume is identified by intersection with two parallel planes (i and j). However, if the ray is close to being parallel with the two planes (ray s in the image), it is not identifiable with this method (left). Rohlfing et al. [56] addressed this problem by suggesting a cylinder to be used instead of the two planes. However, the ray s is still not identifiable if it is close to being parallel with the axis of such a cylinder (center). We suggest the use of the outside surface of the volume for ray identification. Using this method, any possible ray going through the volume is identifiable by its intersections with the surface of the cuboid and no ray will be considered which does not pass through the volume.

a spherical one could solve this problem, however it is too complicated to implement in most medical applications. Also, as long as a sphere is used for identification and sampling of the rays, the problem of inefficient use of memory exists due to the portion of rays which do not cross the volume, but occupy memory in the look-up table.

In this work, we have suggested the use of a cuboid for sampling the rays. This method has the advantage that it supports any ray passing through the volume and therefore eliminates any limit on the transformation of the volume. Moreover, since the cuboid coincides with the surface of the volume, it is guaranteed that no memory or processing time is assigned to rays that do not pass through the volume and could never have an attenuation value.

2.2 Method

Our method is based on the fact that any ray passing through a cuboid could be uniquely identified by its two intersections with the surface of the cuboid (Figure 2.5). This can be verified by the fact that any given cuboid and line have:

1. two intersection points, if the line passes through the cuboid;

2. infinite intersection points, if the line is tangent to one of cuboid faces;

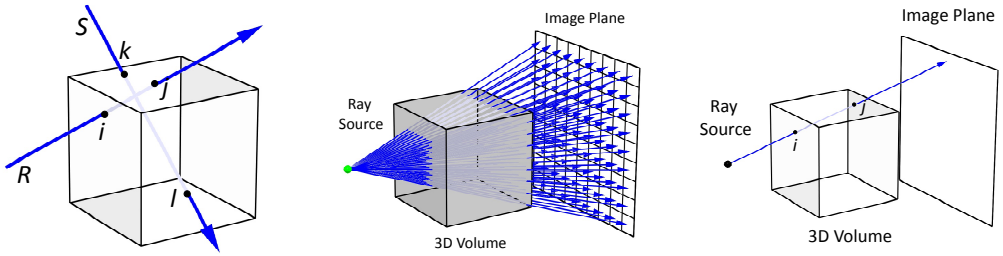


Figure 2.5: The attenuation box method is based on the idea that any line passing through a cuboid could be identified with its two, and only two, intersection points with the surface of the cube (left). All of the rays needed during DRR generation with attenuation values other than zero have to have gone through the volume (middle). During DRR generation, each ray is identified using its intersection points with the surface of the volume and its attenuation value is read from the precalculated look-up table.

3. no intersection points if the line does not touch the cuboid.

Any other number of intersection points between a cuboid and a line are infeasible due to the convex nature of the cuboid.

The implication of the above paragraph is that any ray with an attenuation value other than zero could be identified with its two, and only two, intersection points with the surface of the 3D volume. This provides us with a comprehensive set of all rays, the attenuation value of which we might possibly need during DRR generation. This set is all possible pairs of points on the surface of the volume. Therefore, if the attenuation values of any ray in this set is precalculated, it will be possible to generate any possible DRR of the volume with only reading the precalculated ray attenuation values pertaining to each pixel in the DRR. Such a method would support any position of the volume in the 3D space, because if a ray passes through the volume, its attenuation value would have been precalculated, and if the ray does not pass through the volume or is tangent to its surface, the attenuation value is zero.

2.2.1 Precalculation of ray attenuation values

To precalculate the ray attenuation values, a sampling strategy is required to sample the surface points of the volume. This is because such points form a continuous set, which results in infinitely many rays to preprocess. Therefore, each side of the volume is sampled according to a rectangular grid (Figure 2.6). The distance between two adjacent parallel grid lines will be

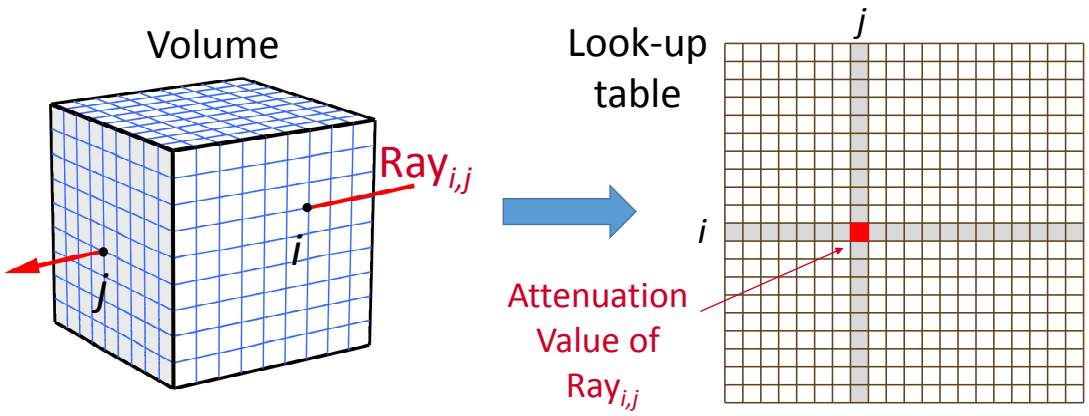


Figure 2.6: Ray passing through the volume are sampled using their intersection points with the surface of the volume on a rectangular grid. The attenuation value of each sampled ray is then calculated and stored in a lookup table in the form of a symmetric matrix and is retrieved during DRR generation.

called ray sampling distance, ρ, hereafter. The smaller this distance, the higher the resolution of the sampling of rays, which leads to a higher quality in the resulting DRRs.

Since the sampling of all of the points on the surface of the volume involves six different sampling grids - one for each cuboid side - a mapping system is needed to map any sampled point to a unique linear index. Such an algorithm is provided in algorithm 1. This algorithm maps each point, $A = (a_1, a_2, a_3)$ of the grid defined on the surface of the cube to a unique linear integer in the range $\{0, 1, ..., \gamma\}$, in which $\gamma = 2 * (g_1 * g_2 + g_1 * g_3 + g_2 * g_3)$ is the number of sampled surface points and g_1, g_2, and g_3 are the number of grid lines perpendicular to each side of the cuboid. Hereafter, G will be referred to as the size of the sampling grid, where $G = (g_1, g_2, g_3)$. All calculations are done in the 3D image coordinate system. Each condition in algorithm 1 determines which side of the cuboid the input grid point is on and maps the two dimension grid point to a linear index in the corresponding partition of the index array. This algorithm, not only maps the grid points to their corresponding index, but also rounds any surface point to the closest grid point if the input point is not a grid point.

The reverse of what algorithm 1 does is performed by algorithm 2: it takes the index, idx, of any grid point, $P = (p_1, p_2, p_3)$, as an input and returns the point in the coordinate system of the 3D image. Similar to the aforementioned algorithm, each condition in this algorithm

Algorithm 1 Surface point to index mapper

if $a_3 = 0$ **then**
    $idx \leftarrow a_1 + a_2 * g_1$
else if $a_3 = g_3 - 1$ **then**
    $idx \leftarrow a_1 + a_2 * g_1 + g_1 * g_2$
else if $a_2 = 0$ **then**
    return $idx \leftarrow a_1 + a_3 * g_1 + 2g_1 * g_2$
else if $a_2 = g_2 - 1$ **then**
    $idx \leftarrow a_1 + a_3 * g_1 + 2g_1 * g_2 + g_1 * g_3$
else if $a_1 = 0$ **then**
    $idx \leftarrow a_2 + a_3 * g_2 + 2(g_1 * g_2 + g_1 * g_3)$
else
    $idx \leftarrow a_2 + a_3 * g_2 + 2(g_1 * g_2 + g_1 * g_3) + g_2 * g_3$
end if

determines which side of the cuboid the point belongs to the grid of. In the body of the condition, the linear index is transformed to the corresponding point. Algorithm 2 is used in the preprocessing phase, while algorithm 1 is used during DRR generation.

In the preprocessing step, the attenuation value of all rays formed by the grid points is calculated and stored. All such rays are shown as R_{ij}, in which i and j are the indexes of the points at which the ray enters and exits the cube respectively. In the preprocessing phase, for any combination of indexes, (i, j), the corresponding ray is formed by connecting the grid points and the attenuation value of the ray R_{ij} is calculated. This calculation is performed by a sampling of the volume intensities at equal distances across the line segment connecting the grid points i and j (Figure 2.6). The intensity of each sampled point is obtained by a trilinear interpolation of the intensities from the eight closest voxels. In this work, the average intensity of the sampled points on the line segment was set as the attenuation value of the ray. However, any other attenuation formula could be used instead.

The attenuation value of the rays is then stored in a matrix, T, in which t_{ij} is the precalculated attenuation value of R_{ij}. This matrix will be called the look-up table hereafter. Since the intensity attenuation of R_{ij} is always equal to that of R_{ji}, the look-up table will be in the form of a symmetric matrix. The look-up table can, then, be stored in an array, V, so that $t_{ij} = V_{i(i+1)/2}$. Using this data structure, only $G(G+1)/2$, of the ray intensity attenuation values need to be calculated and saved, resulting in the required memory and calculation to be reduced in half.

Algorithm 2 Index to surface point mapper

if $idx < g_1 * g_2$ **then**
    $p_1 \leftarrow idx \mod g_1$
    $p_2 \leftarrow \lfloor idx/g_1 \rfloor$
    $p_3 \leftarrow 0$
else if $idx < 2g_1 * g_2$ **then**
    $p_1 \leftarrow (idx - g_1 * g_2) \mod g_1$
    $p_2 \leftarrow \lfloor (idx - g_1 * g_2)/g_1 \rfloor$
    $p_3 \leftarrow g_3 - 1$
else if $idx < 2g_1 * g_2 + g_1 * g_3$ **then**
    $p_1 \leftarrow (idx - 2g_1 * g_2) \mod g_1$
    $p_2 \leftarrow 0$
    $p_3 \leftarrow \lfloor (idx - 2g_1 * g_2)/g_1 \rfloor$
else if $idx < 2(g_1 * g_2 + g_1 * g_3)$ **then**
    $p_1 \leftarrow (idx - 2g_1 * g_2 - g_1 * g_3) \mod g_1$
    $p_2 \leftarrow g_2 - 1$
    $p_3 \leftarrow \lfloor (idx - 2g_1 * g_2 - g_1 * g_3)/g_1 \rfloor$
else if $idx < 2(g_1 * g_2 + g_1 * g_3) + g_2 * g_3$ **then**
    $p_1 \leftarrow 0$
    $p_2 \leftarrow (idx - 2(g_1 * g_2 + g_1 * g_3)) \mod g_2$
    $p_3 \leftarrow \lfloor idx/g_2 \rfloor$
else
    $p_1 \leftarrow g_1 - 1$
    $p_2 \leftarrow (idx - (2(g_1 * g_2 + g_1 * g_3) + g_2 * g_3)) \mod g_2$
    $p_3 \leftarrow \lfloor (idx - (2(g_1 * g_2 + g_1 * g_3) + g_2 * g_3))/g_2 \rfloor$
end if
$p_i \leftarrow p_i *$ sampling resolution , for $i = 1, 2, 3$

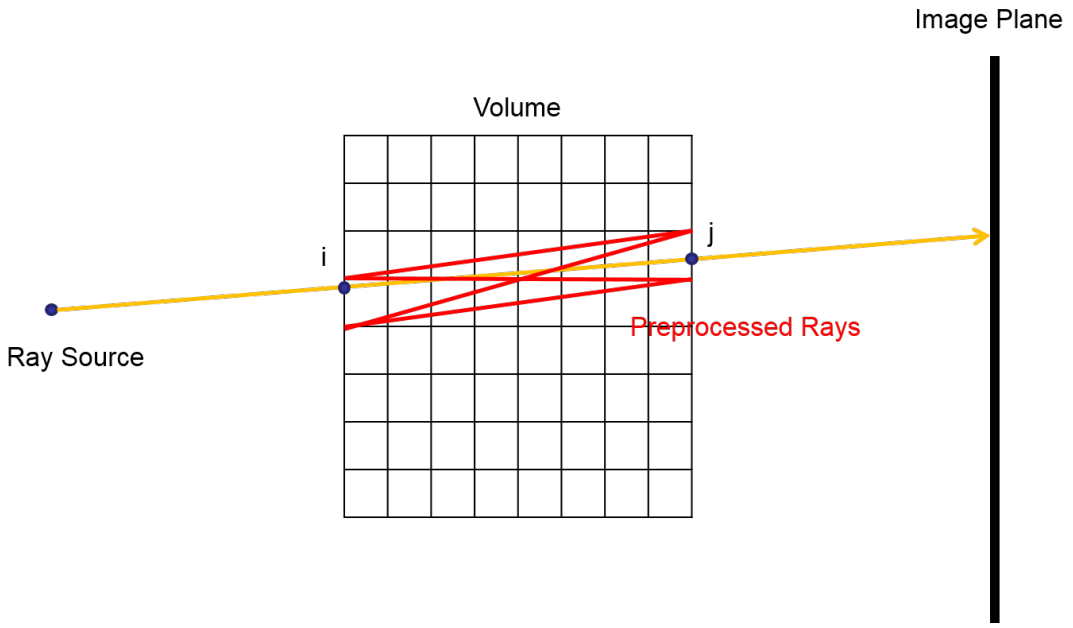


Figure 2.7: When generating a DRR, instead of calculating the attenuation value of each ray, this value is approximated by applying quadrilinear interpolation on the sixteen closest sampled rays. This is done by find the intersection points of the ray and the surface of the volume and finding the four closes surface sampling points to each of them.

2.2.2 DRR generation

The generation of any DRR is possible with the precalculated look-up table. As in the ray-casting method, for every pixel in the DRR, the ray from the ray source to the that pixel is formed and the two intersections, if any, of the ray and the volume surface is found. However, in the attenuation box method, instead of sampling the volume intensities along the ray, the precalculated attenuation value of the ray is read from the look-up table and stored as the intensity of the current DRR pixel. To achieve this, the indexes of the closest grid points to the intersections is calculated using algorithm 1. Thereafter, the attenuation value t_{ij} is looked-up in T, where i and j are the obtained grid point indexes.

2.2.3 Calculating the look-up table

Since the intensity attenuation of ray R_{ij} is always equal to that of R_{ji}, the look-up table will be in the form of a symmetric matrix. The look-up table can, then, be stored in an array, V, so that $L_{i,j} = V_{i(i+1)/2}$. Using this data structure, only $G(G+1)/2$, of the ray intensity attenuation

values need to be calculated and stored. This causes the required memory and calculation to be reduced by half.

Before any DRR can be generated, the attenuation value for each ray needs to be calculated and stored in the look-up table. For any ray R_{ij}, with $i, j \in \{0, 1, \ldots, G-1\}$, the surface points corresponding to indices i, and j are calculated using algorithm 2. Then, the line segment connecting the two points is found and the intensity of the points on the line segments are sampled at equal distances. The intensity of each point is obtained by performing the tri-linear interpolation of the adjacent voxels. Thereafter, the average intensity of these points was saved at location (i, j) of the look-up table.

2.3 Experiments

The experiments were performed on volumes of three bone models: C2, C3, and ulna. Bones were reconstructed from CT image series of human cadaver. The reconstruction was done in ITK-Snap [76] and the size of the volumes was $67.52 \times 46.49 \times 38.75$ mm ($167 \times 115 \times 62$ pixels of size $0.404 \times 0.404 \times 0.625$ mm) for human ulna, ($222 \times 165 \times 38$ pixels of size $0.26 \times 0.26 \times 0.625$ mm) for C2 vertebra and ($232 \times 168 \times 37$ pixels of size $0.26 \times 0.26 \times 0.625$ mm) for C3 vertebra. For each volume different look-up tables were generated at different ray sampling resolutions. The resolutions were $\{0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0\}$ millimeters to study the behaviour of the algorithm in fine resolutions and $\{3, 4, 5, 6, 7, 8\}$ millimeters for lower resolutions (Figure 2.8). Thereafter, for each of these libraries, 200 DRR's were produced with the bone volume set at random positions in terms of rotation and translation. The translation parameters were constrained so that the totality of the bone would stay within the DRR image. The quality of the produced DRR's was then compared to DRR's generated with identical bone position and camera parameters, but through the traditional ray-casting method. Without loss of generality, the intensity value of each pixel was between 0 and 255 in the DRR's generated from either method. Therefore, the memory for storing the intensity value of each ray was one byte.

The quality of the images was measured using two image similarity measures: the ratio of unchanged pixels (RP) [32], and the peak signal-to-noise ratio (PSNR) [30]. The first measure

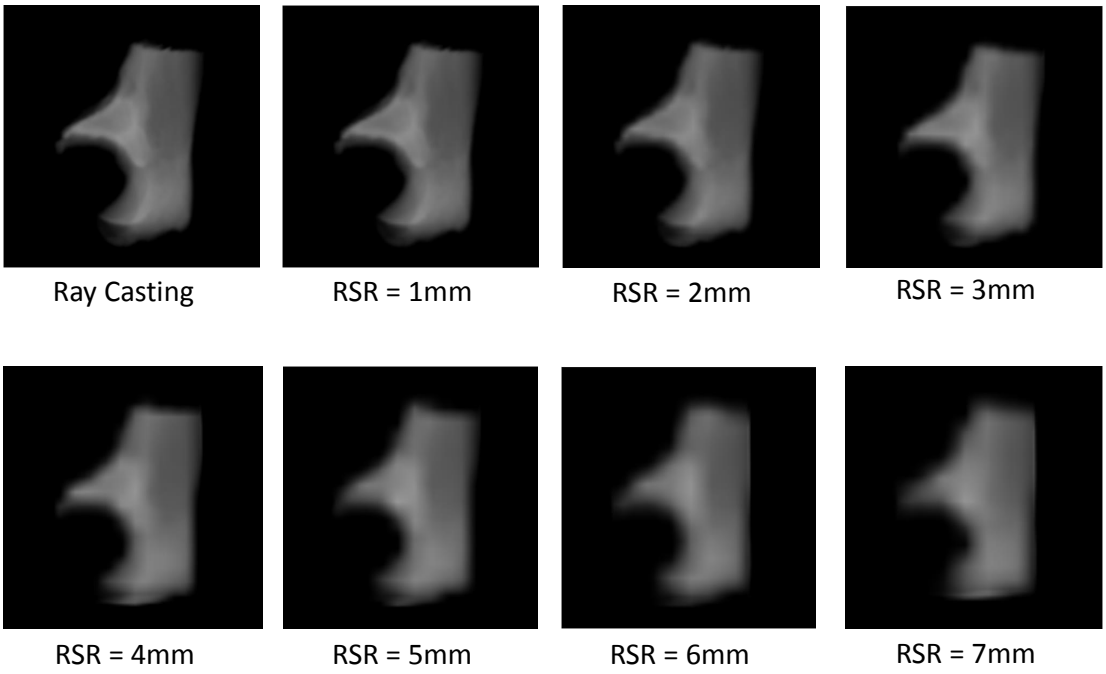


Figure 2.8: A sample of generated DRRs for human ulna at different ray sampling resolution. Any ray sampling resolution of higher than 2 millimeters resulted in high quality images. At a ray sampling resulion of 1 millimeters and above, there is no visible difference between the DRRs generated from the ray casting and those from the attenuation box method.

is extracted from a reliability measure known as kappa statistic [4] and is calculated as

$$RP = \frac{2|A \cap B|}{|A| + |B|}$$

in which A and B are the set of non-zero pixels in the two images. The PSNR, on the other hand, is an estimation of human perception of reconstruction quality. It is measured in the logarithmic decibel scale (dB) and is calculated by

$$PSNR = 20log_{10}\frac{S}{N}$$

. A PSNR value more than 36 dB is generally conceived to represent excellent image quality [30, 58].

2.4 Results

Tables 2.1, 2.2, and 2.3 show the details of the experiment results for every ray sampling resolution. For every produced library, the average PSNR and RP, along the 200 produced random DRRs could be seen in the first and second columns. With all the bone volumes studied, the produced DRRs have been able to reach a PSNR of 36 dB with a sampling resolution as low as 2 mm. At this resolution, 96% of the pixels in the DRRs resulted from the AB method have the same intensity as the ones from the ray casting method. Lowering the ray sampling resolution results in further increase in the quality of the generated DRRs up to 49.51 dB at a ray sampling resolution of 0.6 millimeters. At this resolution, 99.5% of the pixels share identical intensities as those in the ray casting method.

DRR generation using the AB method showed to be approximately 75 times faster than the ray casting method for the ulna bone and 50 times faster for the vertebrae, which have a smaller size compared to the ulna. The longest average DRR generation time was 73 milliseconds for ulna and 56 milliseconds for the vertebrae. Whereas, DRR generation time for ray casting was around 5.5 seconds for the ulna and 2.5 seconds for the vertebrae bone volumes.

DRR generation time in the AB method showed independence from image quality. Higher DRR quality, however, came at the price of prolonged preprocessing time and increased memory. DRR generation time remain within a span of 20 milliseconds for each bone regardless of ray sampling resolution and image quality. The shortest time and memory requirements for high quality DRR generation was 99 and 30 seconds and 7.45 and 3.15 megabytes for the ulna and the vertebrae respectively. The maximum requirements for the highest level of quality produced in our studies was 3 hours 819 megabytes for the ulna and 54 minutes 354 megabytes for the vertebrae.

2.5 Discussion

In this chapter, we have proposed a fast method for the calculation of DRRs through preprocessing of the ray attenuation values. The generation of DRRs is a time-intensive problem of complexity $O(n^3)$, which is the speed limiting step in many medical applications including the registration of 3D to 2D images. The idea of preprocessing a sample of possible ray values

Table 2.1: The experiment results for attenuation box DRR generation and preprocessing at different ray sampling resolutions. The attenuation (AB) box shows a two order of magnitude decrease in DRR generation time compared to ray casting (RC). Due to the larger size of the Ulna volume compared to the human vertebra, the reduction in DRR generation time is more significant for the Ulna. This is due to the fact that AB DRR generation time depends on the size of the DRR only and is independent of the volume size. In terms of quality, a peak signal to noise ratio (PSNR) of greater than 36 dB is obtained at a ray sampling resolution of 2mm and above. The preprocessing time at this resolution is 100 seconds and the library size is 7.45 megabytes (MB).

RSR (mm)	PSNR (±STD) (dB)	RP	Preprocessing Time (s)	Library Size (MB)	AB Time (ms)	RC Time (ms)
0.6	49.51 (±0.75)	0.995 (±0.001)	10656.46	818.88	73 (±13)	5564 (±714)
0.7	49.15 (±0.80)	0.995 (±0.000)	05781.72	443.36	72 (±14)	5530 (±707)
0.8	48.45 (±0.93)	0.996 (±0.001)	03459.27	260.57	72 (±13)	5528 (±674)
0.9	47.96 (±0.94)	0.995 (±0.001)	02192.80	167.15	72 (±14)	5604 (±658)
1.0	47.11 (±1.16)	0.995 (±0.001)	01496.08	112.54	74 (±12)	5652 (±722)
1.1	46.20 (±1.11)	0.995 (±0.002)	00987.87	074.37	72 (±12)	5475 (±639)
1.2	45.29 (±1.09)	0.994 (±0.001)	00682.49	052.06	71 (±13)	5434 (±679)
1.3	44.83 (±1.06)	0.993 (±0.002)	00511.14	038.83	72 (±12)	5656 (±640)
1.4	43.83 (±1.28)	0.992 (±0.002)	00374.09	028.30	71 (±13)	5373 (±600)
1.5	42.82 (±1.21)	0.991 (±0.002)	00286.66	021.68	71 (±13)	5446 (±704)
1.6	42.17 (±1.24)	0.990 (±0.002)	00211.66	016.29	70 (±12)	5240 (±643)
1.7	41.75 (±1.28)	0.989 (±0.002)	00182.17	013.74	71 (±12)	5534 (±674)
1.8	41.13 (±1.32)	0.988 (±0.002)	00146.03	010.95	72 (±14)	5455 (±712)
1.9	40.54 (±1.42)	0.986 (±0.003)	00124.62	009.07	72 (±12)	5585 (±661)
2.0	39.94 (±1.30)	0.985 (±0.003)	00099.48	007.45	63 (±11)	5527 (±730)
3.0	34.72 (±1.41)	0.971 (±0.005)	00019.53	001.46	69 (±12)	5151 (±642)
4.0	31.23 (±1.47)	0.956 (±0.010)	00006.27	000.47	58 (±10)	4911 (±586)
5.0	29.10 (±1.23)	0.941 (±0.012)	00002.98	000.21	67 (±12)	5049 (±582)
6.0	27.19 (±1.42)	0.929 (±0.015)	00001.59	000.11	67 (±12)	4852 (±661)
7.0	24.86 (±1.55)	0.916 (±0.018)	00000.85	000.06	64 (±12)	4581 (±581)
8.0	23.33 (±1.38)	0.907 (±0.024)	00000.51	000.03	52 (±09)	4039 (±553)

Table 2.2: Attenuation preprocessing and DRR generation results for human C2 vertebra at different ray sampling resolutions. The DRR generation using AB is roughly fifty times faster than the ray casting method. Similar to the results from the human Ulna vone, the method produced high quality (PSNR¿36dB) images at any ray sampling resolution higher than 2 mm. The minimum preprocessing time and memory requirements for generation high quality images was 26.6 seconds and 2.79 megabytes.

RSR (mm)	PSNR (±STD) (dB)	RP	Preprocessing Time (s)	Library Size (MB)	AB Time (ms)	RC Time (ms)
0.6	47.83 (±0.69)	0.991 (±0.002)	02947.90	326.58	55 (±13)	2411 (±275)
0.7	46.99 (±0.88)	0.991 (±0.002)	01616.03	177.77	55 (±11)	2439 (±267)
0.8	45.77 (±0.96)	0.991 (±0.002)	00936.69	103.30	53 (±12)	2356 (±259)
0.9	45.08 (±0.97)	0.990 (±0.002)	00586.16	064.63	53 (±11)	2346 (±252)
1.0	44.11 (±1.25)	0.989 (±0.003)	00393.24	043.40	55 (±12)	2397 (±260)
1.1	43.37 (±1.19)	0.987 (±0.003)	00293.44	030.98	54 (±11)	2419 (±271)
1.2	42.11 (±1.29)	0.985 (±0.004)	00196.06	021.28	53 (±11)	2396 (±273)
1.3	41.16 (±1.36)	0.983 (±0.005)	00140.16	015.36	52 (±11)	2300 (±237)
1.4	40.58 (±1.39)	0.982 (±0.005)	00107.06	011.33	52 (±11)	2297 (±261)
1.5	39.84 (±1.46)	0.979 (±0.006)	00081.06	008.57	52 (±10)	2287 (±240)
1.6	39.10 (±1.52)	0.977 (±0.007)	00066.05	007.01	54 (±11)	2356 (±259)
1.7	38.42 (±1.49)	0.974 (±0.007)	00051.49	005.36	52 (±11)	2290 (±261)
1.8	37.84 (±1.53)	0.972 (±0.008)	00041.30	004.27	51 (±11)	2219 (±250)
1.9	37.27 (±1.52)	0.969 (±0.008)	00035.29	003.63	53 (±11)	2371 (±270)
2.0	36.54 (±1.49)	0.967 (±0.009)	00026.61	002.79	45 (±08)	2227 (±234)
3.0	31.76 (±1.37)	0.939 (±0.015)	00005.24	000.54	49 (±09)	2013 (±207)
4.0	29.32 (±1.35)	0.917 (±0.018)	00002.02	000.20	43 (±09)	2029 (±227)
5.0	27.26 (±1.55)	0.897 (±0.024)	00000.96	000.09	48 (±11)	1982 (±214)
6.0	24.86 (±1.23)	0.872 (±0.030)	00000.40	000.04	44 (±09)	1602 (±167)
7.0	24.42 (±1.86)	0.858 (±0.039)	00000.29	000.02	47 (±09)	1894 (±204)
8.0	22.55 (±1.52)	0.817 (±0.045)	00000.19	000.02	41 (±10)	1675 (±188)

Table 2.3: Preprocessing and DRR generation results using attenuation box for human C3 vertebra. The results are similar to that of the C2 vertebra, which is due to the similarity of the volumes in terms of size. The algorithm consistently generates high quality DRRs (PSNR¿36dB) at ray sampling resolutions of 2 millimeters and above. The minimum preprocessing time and memory size for high quality DRR generation was 30.6 seconds and 3.2 megabytes.

RSR (mm)	PSNR (±STD) (dB)	RP	Preprocessing Time (s)	Library Size (MB)	AB Time (ms)	RC Time (ms)
0.6	47.68 (±0.74)	0.991 (±0.002)	03274.16	354.15	56 (±11)	2467 (±279)
0.7	46.67 (±0.98)	0.990 (±0.002)	01827.84	195.39	55 (±12)	2526 (±297)
0.8	45.72 (±1.11)	0.990 (±0.002)	01055.25	113.33	54 (±12)	2472 (±316)
0.9	22.61 (±5.42)	0.789 (±0.124)	00669.87	071.72	54 (±11)	2401 (±285)
1.0	43.84 (±1.47)	0.989 (±0.003)	00436.07	046.72	55 (±12)	2392 (±273)
1.1	42.95 (±1.42)	0.987 (±0.004)	00292.52	031.43	55 (±11)	2431 (±287)
1.2	42.00 (±1.38)	0.986 (±0.004)	00210.57	022.50	54 (±12)	2373 (±287)
1.3	41.07 (±1.59)	0.984 (±0.005)	00153.94	016.49	54 (±11)	2394 (±267)
1.4	40.41 (±1.51)	0.982 (±0.006)	00123.80	012.81	53 (±11)	2481 (±254)
1.6	38.88 (±1.60)	0.978 (±0.006)	00069.39	007.32	54 (±10)	2439 (±302)
1.7	38.25 (±1.54)	0.977 (±0.007)	00057.87	005.94	55 (±12)	2430 (±287)
1.8	37.30 (±1.61)	0.973 (±0.008)	00042.62	004.48	54 (±11)	2291 (±263)
1.9	36.64 (±1.60)	0.971 (±0.009)	00035.88	003.72	54 (±12)	2348 (±289)
2.0	36.27 (±1.55)	0.968 (±0.010)	00030.55	003.15	47 (±10)	2435 (±298)
3.0	31.37 (±1.44)	0.945 (±0.016)	00006.49	000.64	51 (±11)	2254 (±272)
4.0	28.03 (±1.41)	0.919 (±0.022)	00001.97	000.20	43 (±08)	2039 (±230)
5.0	26.09 (±1.47)	0.896 (±0.032)	00000.93	000.09	50 (±11)	1988 (±224)
6.0	24.36 (±1.34)	0.864 (±0.044)	00000.48	000.04	48 (±11)	1831 (±214)
7.0	23.95 (±1.72)	0.844 (±0.052)	00000.38	000.03	51 (±11)	2218 (±260)

has been employed in the attenuation field method [58, 56] to reduce the complexity of the algorithm. However, the range of supported spatial transformations for the volume is limited for each library of rays and using multiple libraries consumes a impractically large amounts of memory adding to the complexity of the method.

In this work, we have proposed an improvement over the attenuation field method by sampling all possible ray attenuation values prior to DRR generation. This is possible using a sampling grid on the surface of the volume and a mapping system, which maps each grid to a linear index. Thus the library of the precalculated ray attenuation values is stored in the form of a two dimensional matrix, which is fast to access during DRR generation. This library includes a sample of all possible rays passing through the volume, which is sufficient for generating DRRs regardless of the (translational and rotational) transformation of the volume. At the time the attenuation value of a ray is needed, this value is calculated by performing quadrilinear interpolation on the sixteen closest sampled rays.

Our proposed method, namely attenuation box, reduces the complexity of DRR from $O(n^3)$ to $O(n^2)$ due to the fact that a sampling of voxels for each ray attenuation value during DRR generation is no longer needed. In our experiments, the attenuation box method showed a reduction of 50 to 40 times in DRR generation time. The experiments were done on volumes of human ulna and C2, and C3 vertebrae. The reduction in DRR generation time was larger for the human ulna, which was due to the larger size the volume. Since, unlike the traditional ray casting method, the complexity of the AB method is independent of the size of the volume, the reduction in DRR generation time is more significant for larger volumes.

The attenuation box method showed to achieve high quality results, on par with the tradition ray casting method, within reasonable time and memory requirements. The quality of the generated DRRs is a function of the ray sampling resolution, which is decided by the user. Across all of the volumes, any ray sampling resolution of higher than 2 millimeters lead to DRRs with PSNRs of higher than 32 dB, in which 98.5 percent of non-zero pixels have the same value as the DRRs from the ray-casting method. The required memory for such resolution was 7.45 MB. The highest quality DRRs in our experiments were generated using the highest ray sampling resolution. At a ray sampling resolution of 0.6 millimeters the PSNR of the generated DRRs was 49.5 dB, in which 99.5% of the non-zero pixels had the same intensities as in the

DRRs generated using the ray casting method. Although an even further improvement of the image quality is possible, at this resolution the DRRs generated using either either method were practically of the same quality.

DRR generation time in the AB method showed to be independent of the image quality. This means a finer quality in the DRRs does not come at the price of longer registration time, which is due to the fact that DRR image quality depends on the sampling of the rays, which is done in the preprocessing step and prior to DRR generation. This, however, means that a higher quality in the DRRs will involve longer preprocessing time and memory requirements. These requirements showed to be low in our experiments. A preprocessing times as low as 99.5 seconds was able to generate DRRs of high image quality with a library size of 7.5 megabytes. Even an extreme increase in the image quality increased the preprocessing time to less than three hours with a library size of 819 megabytes, which is quite smaller than the RAM sizes available today.

The efficiency of the attenuation box has been demonstrated in our experiments and it has been used for DRR generation all through this thesis. However, the performance of the method on large volumes is still to be investigated. Given the same ray sampling resolution, the number of sampled surface points is a d^3, where d is the size of the volume. The number of the precalculated rays is a square of the number of sampling points on the surface of the volume. This results in a fast growth in preprocessing time and library size as the size of the volume increases. Although, the small preprocessing time and memory in our experiments show that the method will be reliable for larger volumes of individual bones, the method might become too cumbersome for substantially larger DRRs such as who body images. The performance of this method on such volume needs investigation. In case of excessive preprocessing time, the recently common use of graphic processing units (GPU) for parallel computing of the rays [19] could be of interest for the attenuation box. Also, a data compression mechanism, such as in the attenuation field [59], could substantially reduce the size of the library. In the current work, however, the resource requirements of the algorithm are well beneath the threshold for such methods to be beneficial.

2.6 Conclusion

DRRs are simulated radiographic images, which are generated in large numbers in medical 3D/2D image registration algorithms. Due to the high complexity of DRR generation process, such registration algorithms are time-intensive, which hinders their application in the medical field. To avoid the time-intensity, several alternate DRR generation methods have been suggested, which generated faster DRRs at the cost of image quality, or the generality of volume transformation. In this work, we proposed a new method, namely attenuation box, for a faster generation of DRRs without compromising quality or generality. This method works by a preprocessing of the rays based on their intersection points with the outer surface of the volume. Our method is an improvement over the existing preprocessing-based DRR generation methods (AF and PAF) in that it supports any possible rigid transformation of the volume. Also, the only parameter which needs to be decided prior to running the algorithm is the resolution of ray sampling. This is unlike the attenuation fields, where sophisticated geometrical calculations need to be done to determine the specifications of the planes.

In our experiments, the Attenuation Box was able to reduce DRR generation time two orders of magnitude. The DRR generation time was independent of ray sampling resolution, which means increasing the quality of the image will not result in longer processing time. An increase in image quality will, however, require longer preprocessing time and larger memory for the look-up table. In our results, the AB created high quality DRRs ($PSNR > 36dB$) with with ray sampling resolutions as high as 2 mm. This means high quality results were produced with preprocessing time as short as 99 seconds and a look-up table as small as 7.45 megabytes (Table 2.1). Also, at all ray sampling resolutions at or above 1.6 mm, %99 of the pixels in the DRRs generated by the two methods were identical (Figure 3.5). This resolution took a processing time of 4 minutes and a look-up table of size 21.68 megabytes (Table 2.1). These preprocessing requirements are well within the practical limits of computing and enable the AB method to be employed in clinical registration frameworks to reduce the registration time. Future work might include an investigation of the method on larger volumes, parallelization of the preprocessing phase, and using data compression methods on the look-up table.

Chapter 3

Fast Medical 3D/2D Image Registration using Kernel Estimated Probability Density Function of Image Gradient Images as a Feature

3.1 Introduction

Three-dimensional (3D) to two-dimensional (2D) image registration is essential in many medical applications such as preoperative surgery planing, intra-operative therapy guidance [58], patient placement for radiotherapy planning and treatment verification [9, 25], radiosurgery [48], cranial neurosurgery [41, 63], neurointerventions [35, 37], spinal surgery [40, 69], orthopedic surgery [34, 75], and aortic stenting procedures [69, 11, 50, 58]. The goal of 3D to 2D (3D/2D) registration is to find the rigid transformation parameters which convert the coordinate system of a 3D model of the subject to the real clinical one [20, 73]. Such a transform consists of three translation and three rotation parameters along and around X, Y, and Z axes (6 degrees of freedom). These parameters are calculated by finding the 6 degrees of freedom, which maximize the similarity of the clinical 2D radiographic image of the subject to a 2D projection of the 3D model. The X and Y axes are the coordinates of the 2D image plane (the fluoroscopy video frame), and the Z axis is perpendicular to the image plane.

One application of 3D/2D image registration, which motivated the method introduced in this paper, is the estimation of the 3D kinematics of in-vivo bone movement. An accurate estimation of bone kinematics is crucial in the image guided evaluation of musculoskeletal disorders and surgeries. In this application, the bone movement is captured through a 2D fluoroscopy video. To find the 3D pose of the bone at any instant of the motion, a 3D model of the bone, which is reconstructed by the segmentation of a computed tomography (CT) image series, is registered to every frame of the fluoroscopic video. The results are the six transform parameters for every frame of the video. These parameters transform the 3D bone model to the

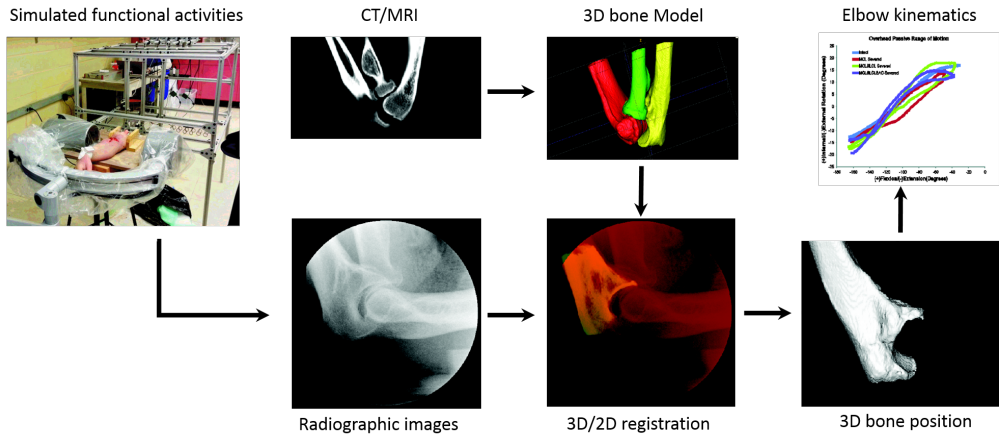


Figure 3.1: Image guided Evaluation of Musculoskeletal Disorder and Surgery through 3D/2D image registration. The C-arm captures the movement of the elbow joint bones through fluoroscopic video frames. The 3D bone models are constructed from CT image series. A model-based tracking approach is used for registering 3D bone models to fluoroscopic images and estimating 3D bone positions. 3D bone kinematics are then calculated from bone positions.

3D bone pose at the time when each of the frames were captured (Fig. 3.1).

The registration of the 3D model to the 2D fluoroscopic image is achieved by matching a 2D perspective projection of the 3D model to the 2D fluoroscopic image. The registration process consists of an optimization algorithm which searches the six dimensional space of all transform parameters for the parameter values which result in the maximum similarity between the 2D fluoroscopic image and a perspective projection of the 3D model onto the image plane. This 2D projection is called the digitally reconstructed radiograph (DRR) and is the result of a simulation of the fluoroscopic imaging process. The camera parameters in the simulation are set up after the camera parameters of the fluoroscopic imaging machine using camera calibration [28]. The DRR is calculated by casting rays from a ray source to an image plane while the rigidly transformed 3D bone model is located in between. The intensity value of each pixel in the DRR is the integral of the volume intensities that the ray from the ray-source to that pixel has passed through [19]. Since the location of the ray-source and the image plane are set up after the fluoroscopy machine, it is assumed that if the DRR matches the fluoroscopy image, the pose of the bone in the 3D space must have been the same as the real bone at the moment the fluoroscopy image is captured.

The registration process is computationally intensive. The registration includes the assessment of many transform parameters, which require the generation of many DRRs. DRR generation is a time-intensive problem due to numerous memory access operations [19]. This problem has usually been addressed by using alternate ways to either estimate or precalculate the DRRs, which come at the cost of quality, or memory and precalculation time [19, 56, 58, 10]. Developing a registration method with the minimal number of required DRRs would be ideal.

In addition to the computational cost, the existing methods are sensitive to the initial guess of the bone position. If the initial position of the bone is too far from the true registration, the optimization algorithm might be stuck in local optima and not converge to registration. Therefore, most existing methods require the user to manually and visually position the 3D model in the vicinity of the optimal answer. A manual location of the 3D model, however, is labor-intensive and prone to error as the 3D position of the bone model, based on its 2D projection is not always evident visually.

In this work, we propose a novel method to address the problems of computation time and algorithm sensitivity to the initial solution. Our method uses a weighted histogram of gradient directions as an image feature to measure the similarity of the DRR and the fluoroscopy image. The histogram works as an estimation of the probability density function (PDF) of gradient directions on the 2D bone image. In the result section, our proposed method is shown to produce high registration accuracy in substantially reduced processing time. This is achieved due to the fact that in our method, the parameters are optimized sequentially rather than simultaneously. In addition to reduction in time, the quality of the results was insensitive to the initial distance from the true registration within 90° of it. This is a substantial improvement over existing methods, which require the 3D volume to be manually registered to the vicinity of the true registration for the method to work.

3.2 Method

The test runs were done on a 3D volumetric model of a ulna bone. The model was reconstructed through a segmentation of a series of CT images from a cadaver elbow. The resolution of the CT images was 0.39 mm by 0.39 mm with an inter-slice distance of 0.62 mm and the segmentation

was performed using the ITK-Snap software [76]. The choice of ITK-Snap for segmentation was arbitrary and our results are independent of the segmentation method. The resolution of the DRRs generated was 250 by 256 pixels with a pixel size of 1.23 by 1.23 millimeters.

The reconstructed ulna model was used in 100 independent runs of our algorithm to test its efficiency. The algorithm was programmed in MATLAB (The Mathworks Inc., Natick, MA, USA). As done in [64], at each run, a DRR resulting from a random position of the bone model was produced and used as a reference (fluoroscopy) image for the rest of the run. To simulate a clinical image, the proximal-distal axis of the bone was set to have an angle less than 45° with the image plane. This position of the bone in space was used to measure the registration accuracy of the algorithm. The difference between the 3D position resulting from the registration and the real one was considered as the registration error. Since the 3D position of the bone is determined by three translation and three rotation parameters, the registration error is shown as six numbers, each depicting the difference between the real value of the relevant transform parameter and the one found by the algorithm.

For the translation and rotation of the 3D bone model, a coordinate system is required. This coordinate system could be defined globally, such as the coordinate system of the DRR image plane, or locally and intrinsic to the bone model. In this work, after the bone model is positioned halfway between the DRR image plane and the ray source, the transformations are done according to the coordinate system of 3D bone model. This coordinate system may be defined clinically, using the anatomical landmarks on the bone, or mathematically.

In our study, the eigenvectors of the bone volume are used as the axes of the bone coordinate system, in which the X and Y axes are the eigenvectors with the largest and smallest eigenvalues (Fig. 3.2). This means the scattering of volume voxels is maximized along X, and minimized along the Z-axis [2]. In other words, it is known that a rotation around the Z-axis could result in the largest change in the 2D projection of the volume and a rotation around the X-axis could result in the smallest. The implication of this fact is that more searching effort should be spent on tuning the X-rotation parameter due to the subtlety of its impact on the DRR. The origin of the coordinate system is set to the center of the volume to minimize the variation in the DRR caused by rotation [43].

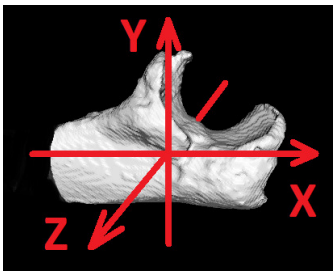
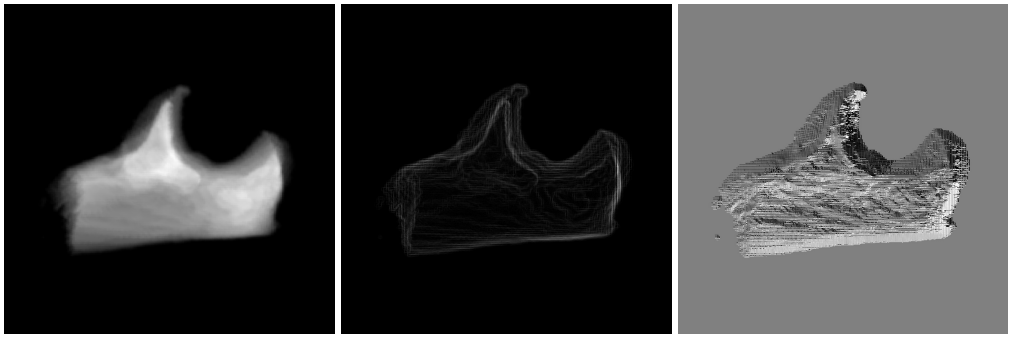


Figure 3.2: The principal components (eigenvectors) of the volume are chosen for establishing the bone coordinate system. The largest principal component is chosen as the X axis and the smallest as the Z axis.

3.2.1 Probability Density Function of Image Gradient Directions

The method proposed in this paper is inspired by three intrinsic properties of the probability density function of gradient directions in a continuous two dimensional Image. A first derivation of intensity values on a gray scale image results in a gradient magnitude and a gradient direction associated with each point on the image [26]. While the magnitude of gradient shows the amplitude of change in intensity, the gradient direction signifies its orientation (Fig. 3.3). An affine transformation of a 2D image translates to specific changes in the probability density function of the image gradient directions, which our method is based on (Fig. 3.4).



(a) Original Image (b) Gradient Magnitude (c) Gradient Direction

Figure 3.3: The magnitude of image gradient (b) shows the velocity of change in intensity at each point/pixel of the image and the gradient direction (c) represents the orientation of the change. This work uses the probability density function of the image gradient directions as a feature for image registration. This function is estimated by a histogram of image gradient directions, which is weighted by the gradient magnitudes.

- The probability density function of gradient directions of an image is insensitive to translation. Since this function is not dependent on the location of the intensities rather the direction of their change, as long as the whole image is preserved, a translation will not change the PDF of gradient directions of an image.

- A rotation of a 2D image is the equivalent of a shift in the probability density function of gradient directions. This is due to the fact that a rotation in the image rotates all directions equally, adding or subtracting a constant to all of the gradient directions without changing the gradient magnitudes. Consequently, the rotation of a 2D image with a certain amount causes only a shift of the same amount in the probability density function of gradient directions of the image.

- A resizing of a 2D image is reflected in the probability density function of gradient directions as a scaling of the function. This is due to the fact that a rescaling of the image causes a decrease or increase of all direction recurrences. The scaling of the function is compensated by an reversely affected probability for direction zero, which is because an enlarging of the image reduces the background area and vice versa.

The implication of these three properties is that if a 2D image is rotated, translated and resized, the re-sizing scale and the rotation angle could be found by first re-scaling the histogram of gradient directions of the transformed image to match that of the original image, and then finding the shift which maximizes the similarity of the two histograms. The amount of this shift will denote the rotation angle. After such a registration, the translation parameters could be found with one of existing methods for transitive image registration. All of the above properties of the probability distribution function of the image gradient are based on the assumption that the image is located in a background so large that the affine transformations do not amputate the totality of the image itself.

In this work, we estimated the probability density function of the gradient directions by applying a Gaussian kernel [2] to the histogram of gradient directions of the image. The direction of gradients in an image shows only the orientation of change in intensity and not its magnitude. Therefore, small changes in the intensity will affect the histogram of gradient directions as much as large ones. To improve accuracy, we have calculated the histogram of gradient

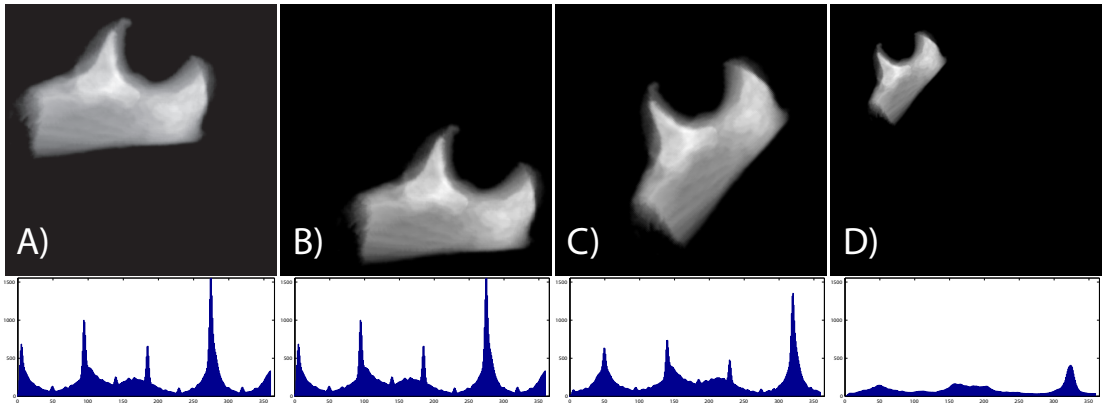


Figure 3.4: Properties of the histogram of gradient directions of a 2D image. a) The original image and its histogram of gradient directions, b) A 2D translation in the image has no effect on the histogram of gradient directions. c) A 45°rotation in the image causes a 45°shift in the histogram of gradient directions. d) A resizing of the rotated image into half its size rescales the height of the bars in the histogram of gradient directions in half.

directions after weighting them by their magnitude. Hereafter, we will call this histogram the weighted histogram of gradient directions (WHGD).

In a perspective 3D to 2D projection, the effect of 3D X-, and Y-translation, Z-rotation, and Z-translation, is similar - but not identical - to 2D X-, and Y-translation, Z-rotation and the scaling of the image respectively. Therefore, the WHGD is nearly insensitive to these transformations. This allows us to register the two images based on their rotation around X and Y axes regardless of their difference in the other parameters.

3.2.2 Distance Measure

In order to extract the WHGD feature from DRR images, the gradient direction and magnitude of every DRR are stored in images GD and GM so that $GD(i,j) \in [0, 360)$ and $GM(i,j) \in [0, 1]$ for every pixel (i, j). The image histogram of GD is then built, in which the gradient direction of each pixel in GD is multiplied by its magnitude in GM. The bins in the histogram span the $[0, 360)$ range into $360/\omega$ equal bins, in which ω is the bin width and experimentally chosen as 1°. The height of the d^{th} bin, $h(d)$, is calculated by the following formula:

$$h(d) = \sum_{(i,j)\in GD_d} GM(i,j), \text{for } d = 0, 1, 2, \ldots, 360/\omega - 1\,,$$

in which $GD_d = \{(i,j) \mid \lfloor GD(i,j)/\omega \rfloor = d\}$. The resulting histogram is then smoothed with a Gaussian kernel of width 10 bins. The resulting histogram is the weighted histogram of gradient directions.

The similarity of the DRR and the fluoroscopy image is measured by comparing their WHGDs. The WHGD is an estimation for the probability distribution of the gradient directions in the image. Therefore, to assess the difference between two WHGDS, the existing distance measures between probability density functions were investigated. Of the existing measures in [14], two measures showed higher experimental performance: city block distance [**?**], and J-divergence distance [39, 31, 61]. These two distance measures are obtained from the following formulas:

$$\text{city block distance} = \sum_{all\ d} |h_1(d) - h_2(d)|\,,$$

$$\text{J-convergence distance} = \sum_{all\ d} |h_1(d) - h_2(d)|.ln\frac{a(d)}{b(d)}\,,$$

in which $a(d) = \max_d\{h_1(d), h_2(d)\}$, and $b(d) = \min_d\{h_1(d), h_2(d)\}$.

Experimenting with the above distance measures showed that their combination is more effective than each of them. Therefore, the summation of the two distance measures was employed as the distance measure, m, in the proposed algorithm:

$$m = \sum_{all\ d} |h_1(d) - h_2(d)| + \sum_{all\ d} |h_1(d) - h_2(d)|.ln\frac{a(d)}{b(d)}\,,$$

$$m = \sum_{all\ d} |h_1(d) - h_2(d)|(1 + ln\frac{a(d)}{b(d)})\,,$$

Since a shift in the WHGD pertains to the rotation of the image around the Z-axis, the distance measure is calculated for every $360/\omega - 1$ possible shift in the WHGD of the DRR and the shift amount resulting in lowest distance is chosen as the rotation parameter around the Z-axis. In other words, calculating the distance measure involves registering the image for Z-rotation.

3.2.3 Optimization

The distance measure is minimized in several stages to find the 3D position of the bone resulting in smallest distance (algorithm 3). The distance measure is first minimized over X- and Y-rotation parameters by performing a sequence of pattern search minimization algorithms (PS) [29, 17]. The PS algorithm starts from an initial solution and improves the distance measure iteratively. Since the PS method is prone to getting stuck in local optima, each time the run is over, the final answer is used as the initial solution for a new run of PS. This continues until there is one run of the PS algorithm with no improvement in the distance measure. The initial mesh size for the pattern search algorithm is 30°and the algorithm stops when the mesh size becomes smaller than 0.05°.

After the distance measure is optimized, the rest of the solution parameters are calculated. Since optimizing the distance measure includes the registration of the rotation parameters, the remaining parameters to be registered are the three translation parameters along X-, Y-, and Z-Axes.

The next step is to find the Z-translation parameter. In the absence of other transformations, the effect of Z-translation on the 2D projection of a 3D object is similar to a resizing of the image. Finding the Z-translation parameter is based on the idea, that if the rotation parameters have been found in the previous steps, the Z-translation could be retraced from the resizing ratio of the two images. To do so, the ratio of the two images was calculated as the ratio of the number of pixels with non-zero gradient magnitudes in each image.

The Z-translation parameter is then obtained using the aforementioned scaling factor. Since, by similarity of triangles, the length of the projection of an object is inversely proportional to the distance between the object and the ray source, this distance could be approximated by multiplying it from the current solution by the square root of the area ratio of the moving image to the reference image. Then, the registration for X-Y translation is performed using the phase correlation method [60].

After all transform parameters are calculated, since the new parameters might result in slightly different DRRs, one last run of pattern search algorithm is performed to fine-tune the

final solution. The final mesh size is reduced to 0.005°, which results in a more precise optimization than before. Also, the search is performed within ±5°of the final answer, because the solution must have come close to optimality in the previous searches and a wider search of the solution space is not necessary. When this run of pattern search is over, the other parameters are updated and the resulting solution is output.

Algorithm 3 Pseudo-code for the registration process

```
Sample the X- and Y-rotation parameter space:
for each sampling X- and Y- rotation parameter do
    if best solution improved then
        Register the rest of the parameters.
    end if
end for
loop
    Run pattern search from the best solution with final mesh size of 0.05°
    if best solution improved then
        Register the rest of the transform parameters.
    else
        Break the loop.
    end if
end loop
Limit the search to ±5° of the best solution.
Run pattern search from the best solution with final mesh size of 0.005°
Register the rest of the transform parameters.
```

During the optimization, the values for the transform parameters other than X- and Y-rotation - which are being optimized - are originally set to that of the initial solution. The initial solution is chosen within 30 mm of the real solution in terms of translation, and 90°in terms of rotation. A large difference between the value of these parameters in the initial solution and the true registration might disturb the optimization process. Therefore, a mechanism was developed to improve the quality of the initial solution through a sampling of the solution space at fix distances prior to the main body of optimization.

The sampling distances were experimentally chosen as 10°and 20°for X-, and Y-rotation parameters respectively and were fixed across the experiments. The ratio of the two distances were obtained from the inverse of the ratio of the two largest eigenvalues of the voxels in the 3D image. This is because the eigenvectors of the volume were chosen as the axes of the bone, around which rotation takes place. Since a larger eigenvalue is indicative of larger scattering

of the points around that axis, the rotation of the bone around that axis results in a more subtle change in DRR, therefore, a finer search with a smaller sampling distance is required to find a favorable rotation parameter. Smaller sampling distances result in a finer initial solution, but increased solution time.

During the optimization of X- and Y-rotation parameters, each time an improvement in the best solution achieved is observed, the other parameters are updated using the aforementioned algorithm, and the similarity measure is updated using the new parameters. Thereafter, and until the next improvement in the similarity measure, these updated parameter values are used for the generation of the DRRs.

To verify the registration results from our algorithm, we compared it to a recent successful method for 3D to 2D registration of bone images, namely the weighted edge matching score (WEMS) [42, 65]. In this method, the similarity of the DRRs and the fluoroscopic image is assessed using an edge-based similarity measure. This similarity is maximized using a genetic optimization algorithm framework, which consists of a consecutive production of solution generations and the survival of the fittest. This method was implemented in MATLAB and compared to our method for the investigation of registration accuracy and parameter analysis.

To reduce registration time, we used the DRR generation method introduced in [24], which decreases the DRR generation time by two orders of magnitude. This reduces the registration time significantly, which was essential for the large number test runs in this work. Due to the high complexity of DRR generation, multiple runs of the WEMS method would take an impractically long time using the traditional DRR generation method. To compare the complexity of different 3D/2D registration frameworks the number of generated DRRs is a better indicator than processing time as the preprocessing time depends on the DRR generation method and the hardware used besides the registration method [42, 65].

In addition to the assessment of registration precision, we performed a sensitivity analysis of the algorithm over its parameters. In our experiments, we investigated the effect of the initial distance from the true registration answer, the width of the Gaussian kernel, and the bin width. Moreover, the necessity of the initial solution improvement step of the algorithm was demonstrated.

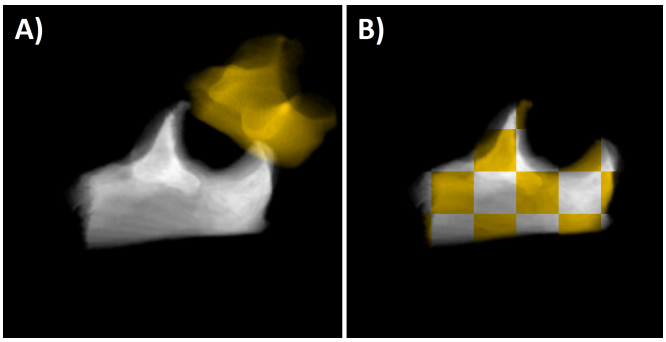


Figure 3.5: Sample registration result for the superior end of the ulna. A) The reference bone position (white) and the initial bone position (yellow). B) The overlapped bone positions after registration.

3.3 Results

In all of the 100 runs, the registration was successful (Fig. 3.5). The mean (±SD) registration error over all of the runs was -0.62 (±0.19), -0.63 (±0.19), and 0.00 (±0.33) millimeters for translation and -0.11°(±0.8°), -0.16°(±0.9°), and -0.04°(±.31°) for rotation along and around X, Y, and Z axes respectively (Table 3.1). In all of the runs, the distance between the initial solution and the true registration was 89° in rotation and 30 mm around and along all of the axes. The average (±SD) number of generated DRRs for one run was 274.76 (±201.73) and the solution time was 12.55 (±10.18) seconds.

To compare the efficiency of our method to WEMS, the population size and the number of generations of the genetic algorithm in WEMS method was experimentally set to 100 and 50 to produce similar accuracy to that of our method. An increase of the generations number beyond 50 showed no improvement in the results as the algorithm consistently converged before the 50th generation. Each run of WEMS required the generation of 5000 DRRs, which took 131.4 (±4.3) seconds of processing time on average (±SD). The results can be seen in table 3.1.

As explained in the method section, the algorithm includes a primary sampling of the solution space to improve the quality of the initial solution. To verify the necessity of this step, one hundred runs of our algorithm was performed without the sampling. The results showed a deterioration of results down to a mean (±SD) absolute registration error of 6.54 (±27.73), 13.89 (±44.8), and 13.7 (±43.47) degrees in rotation and 1.75 (±4.58), 1.83 (±4.51), and 2.13

Table 3.1: Mean absolute (MAE), mean (ME), and standard deviation (SDE) of registration error for WHGD (this paper) and WEMS methods. The initial distance from the true registration was 30 mm and 85°for WHGD method, while this distance was 15 mm and 10°for WEMS. Despite the reduced initial distance, the WEMS method required 5000 DRRs in order to produce similar accuracy to WHGD, whereas this number was 274.76 (±4.3) for WHGD.

		Rotation (°)			Translation (mm)		
		X	Y	Z	X	Y	Z
WHGD	MAE	0.61	0.57	0.27	0.62	0.63	0.26
	SDE	0.80	0.90	0.31	0.19	0.19	0.33
	ME	-0.11	-0.16	-0.04	-0.62	-0.63	0
WEMS	MAE	0.54	0.84	0.31	0.27	0.29	2.07
	SDE	1.3	1.7	0.78	0.61	0.66	3.61
	ME	-0.11	-0.04	-0.13	0.018	0.04	-1.77

(±5.5) millimeters in translation. The initial solution was generated within 30 mm, and 85° of the true registration. The deterioration of the results was less significant with a smaller range of possible initial solutions.

A change in the initial rotational solution distance from the true registration within 90°(180°around the true registration) shows not to affect the accuracy of the results. The algorithm was run with initial solutions generated at 15°, 25°, 45°, 65°, and 85°from the true registration. At each initial range, one hundred runs were performed and the initial translational distance from the true registration was 30 millimeters. In the registration results, the average (±SD) rotational error was -0.06 (±0.69), 0.08 (±0.85), 0.06 (±0.95), 0.09 (±0.99), -0.10 (±0.67) degrees at the aforementioned ranges respectively. The corresponding translational error was -0.39 (±0.25), -0.33 (±1.04), -0.39 (±0.29), -0.3 (±1.05), and -0.42 (±0.24) millimeters. Neither of the error groups showed a correlation with the distance of the initial solution from the true registration.

Although an increase in the initial solution range prolongs the solution time due to a larger initial sampling of the solution space, the final solution time did not show dependence on the initial range as a main factor ($\alpha = 0.05$). This can be seen in Figure 3.7. The average (±SD) number of the generated DRRs in the aforementioned experiments was 120.66 (±124.25), 253.4 (±926), 238.41 (±632), 305 (±691), 274.76 (±201.73), which pertain to solutions times of 6.99 (±7.4), 14.23 (±52.87), 12.47 (±34.42), 15.21 (±36.14), 12.55 (±10.18) seconds for initial ranges of 15°, 25°, 45°, 65°, and 85°. Although the smallest initial range shows a smaller

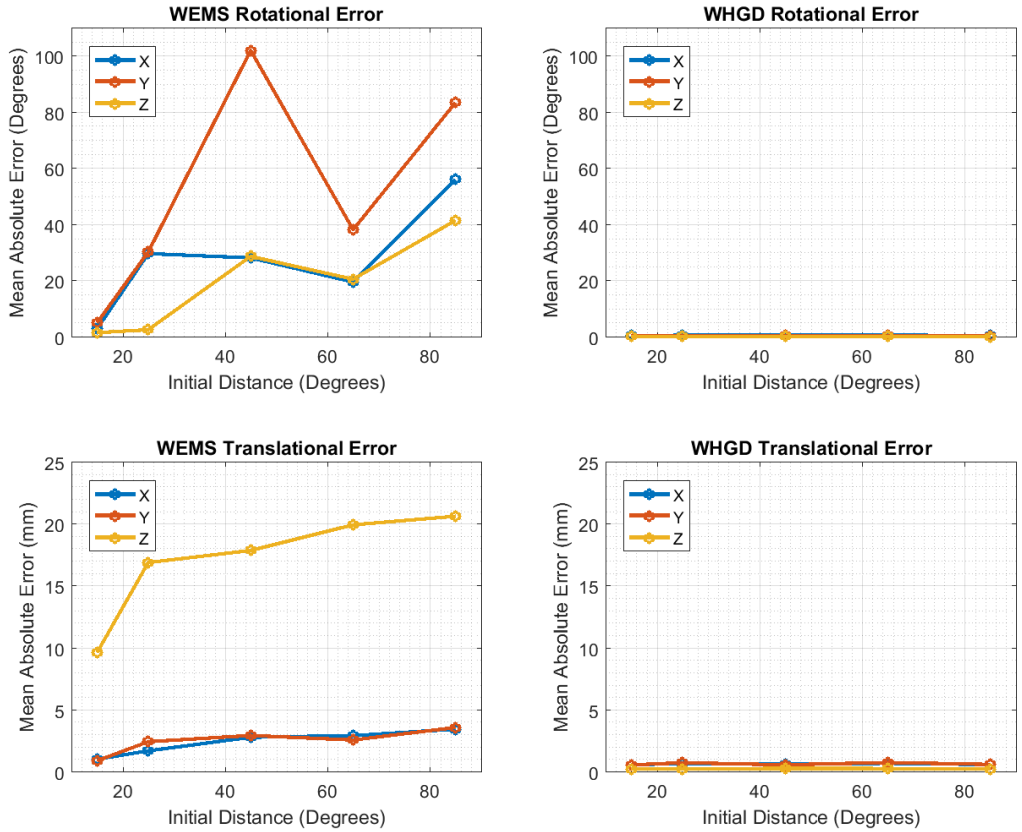


Figure 3.6: The registration results with extended initial distance from the true registration. An increase of initial distance to 85°did not cause any significant change in the quality of the registration in WHGD method, while a substantial deterioration was observed in the WEMS method.

solution time and number of DRRs, the rest of the experiments fail to demonstrate a consistent correlation between the initial distance from the true registration and the solution time.

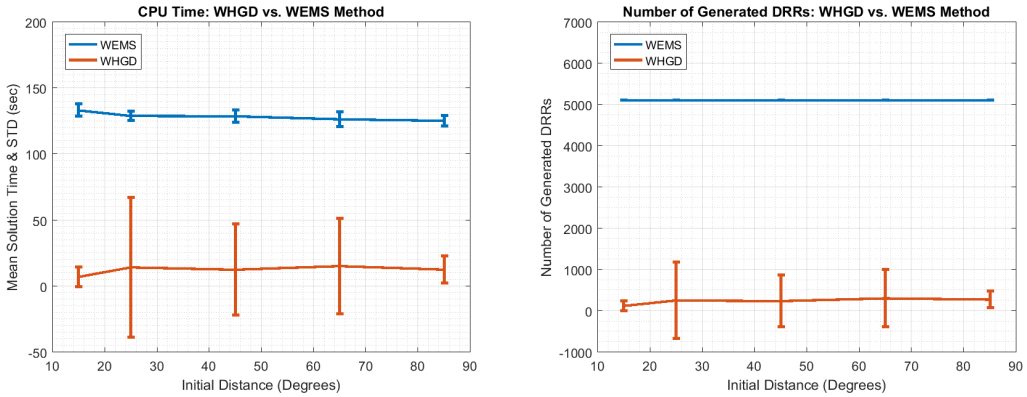


Figure 3.7: Computation time shows significant reduction in WHGD compared to WEMS. The average solution time is smaller in WHGD by one order of magnitude (left). And, the average number of generated DRRs in WHGD is around 5% of the same number in WEMS (right). The vertical lines show the standard deviations for the mean values. The WEMS method has a smaller variation due to the fact that the genetic algorithm uses a fixed number of generations for optimization. Except for an initial distance of 15°, in no other case could the WEMS method achieve the accuracy of the WHGD method.

To compare the robustness of our method to the initial solution, a series of of WEMS algorithms were run at extended initial solution ranges mentioned in the previous paragraph. The average (±SD) error was 0.036 (±4.62), -7.73 (±60.65), -30.29 (±99.05), -4.89 (±45.5), -22.43(±93) degrees rotationally and -3.16 (±3.87), -5.7 (±5.92), -5.69 (±6.48), -6.63 (±5.14), -6.94 (±5.95) millimeters in terms of translation (Fig. 3.6). The number of runs at each initial range was 30 with a population size and generations number of 100 and 50.

The use of the Gaussian kernel for smoothing the histogram of gradient directions showed to improve the accuracy with little sensitivity to the width of the kernel. A set of reruns of the algorithm with kernel widths of 0 to 20 bins (30 runs each) showed a decrease in the mean absolute rotational error from 1.96°at 0 to 0.43°at 10 bins of kernel width. The Z-transnational error also fell from 0.80 mm to 0.26 mm. Increasing the kernel width from 10 to 20 bins does not show a significant change in the accuracy of the results. However, a case of non-convergence was observed at 20 bins and was excluded from the average error (Fig.3.8).

In our experiment results, the width of the histogram bins was inversely proportional to

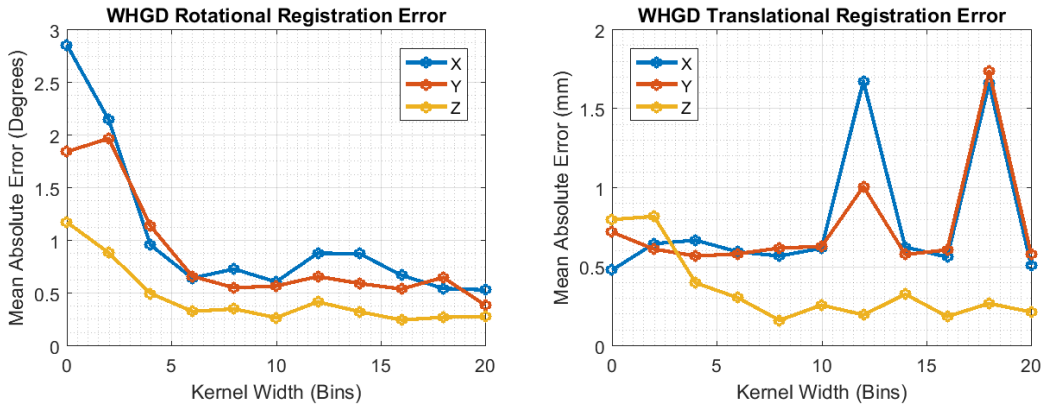


Figure 3.8: The effect of kernel width on registration accuracy in WHGD. The employment of Gaussian kernel smoothing shows improvement in registration results, which increases as the width of the kernel grows from 2 to 6 bins. However, further growth in the width of the kernel does not affect the results until a width of 18 bins. At this width, the algorithm begins to produce cases in which convergence to the answer does not happen. These cases were excluded from the figure.

the accuracy of the results. The mean ($\pm$SD) rotational registration absolute error was 0.48 ($\pm$0.08), 0.59 ($\pm$0.64), 0.68 ($\pm$0.54), 1.21 ($\pm$1.08), and 0.89 ($\pm$0.73) degrees at bin widths of $1, 2, 3, 4, 5$ degrees and 2.66 ($\pm$6.18), 6.54 ($\pm$26.29), 2.53 ($\pm$2.83), 4.79 ($\pm$9.36), and 2.73 ($\pm$2.71) degrees at bin widths of $6, 7, 8, 9, 10$ degrees. A reduction of the bin width to 0.5° resulted in a decrease of the aforementioned error to 0.43° ($\pm$0.54). The number of runs for each bin width was 30 and the initial distance from the solution was 30 mm and 85° (Fig. 3.9).

3.4 Discussion

In this chapter, we proposed a new method for the 3D to 2D registration of images based on the probability density function of the directions of the image gradient. More specifically, the probability density function of the gradient directions is estimated by the histogram of image gradient directions, which is weighted by the gradient magnitudes and smoothed by a Gaussian kernel. The novel contribution of our method is the introduction of the weighted histogram of image gradient directions as an image feature for registration. Such an image feature allows a sequential registration of transformation parameters, which results in substantial decrease in the number of solution assessments.

Figure 3.9: The effect of bin width on registration accuracy. The width of the histogram bins is inversely proportional to the accuracy of the registration results. The deterioration of the results are especially observable for bin widths of greater than 6°. At a bin width of 10°, two cases of non convergence were observed, which are excluded from the figure.

Since the assessment of each solution requires the time-consuming generation of DRRs, this reduction results in shorter computation time for the problem. This method was able to perform the registration not only faster than similar existing methods, but also from a remarkably extended original rotational distance from the answer. Unlike many existing methods, which require an initialization of the volume position close to the correct solution, our method was able to reach the optimum answer from within 90°and 30 mm of the answer in terms of rotation and translation, which is an improvement over the existing methods.

Our method achieves significant increases in both registration speed and tolerance to the initial distance from true registration. We compared the accuracy of the results from our method (WHGD) to that of the weighted edge matching score (WEMS) method introduced by Tsai et al [65] (Fig. 3.1). The WEMS method was not able to produce comparable results to the WHGD method at comparable initial distance from the true registration. Therefore, in order for the WEMS method to generate comparable accuracy to our method, we had to reduce the initial distance to 10° and 15 mm for WEMS, whereas this distance was 85° and 30 mm in WHGD. At this reduced distance, the WEMS method was able to produce comparable results to WHGD. This is while the WHGD method required the generation of only five percent of the DRRs generated by WEMS and the computation time took ten percent of the time for WEMS. The registration error was 0.48° and 0.5 mm for WHGD vs 0.56° and 88 mm for the

WEMS method in terms of mean absolute rotation and translation error. The average number of the generated DRRs was 274.76 in WHGD vs. 5000 in WEMS. The mean computation time was 12.55 seconds in WHGD vs. 125.29 in WEMS. Our method was also able to produce more precise out-of-plane translation (0.26 mm) compared to the WEMS method (2.07) mm. The out of plane translation is a challenging transformation to register due to minimal 2D manifestation of a translation perpendicular to the image plane [65].

In our experiments, the accuracy of registration showed to be independent of the initial distance from the true registration (Fig. 3.6). The average absolute rotational and translational registration error were 0.51° and 48 mm with an initial distance of 15° from the true registration. An increase in this distance from 15° to 85° resulted in an average absolute error of 0.48° and 0.48 mm, which shows no significant change. The translational distance from the true registration was 30 mm in both cases. This is while such an increase in the WEMS method resulted in a practical failure of the registration (Fig. 3.6).

The width of the histogram bins showed to be inversely proportional to the precision of registration (Fig. 3.9). While an increase of the bin width from 1° to 3° showed minimal change in the results, a further growth of the width to 6°resulted in a consistent decline of registration accuracy to a mean absolute rotational error beyond 2° . Increasing the bin width further than 6° resulted in further degradation of the results in addition to sporadic behavior of the algorithm. A bin width of 0.5° showed similar accuracy in our results to a bin width of 1°($\alpha = 0.05$). While the in-plane translation accuracy is dependent on phase correlation and is more a function of the image resolution than algorithm parameters, the out of plane rotation presents similar behavior to the rotations.

In our method, the histogram of gradient directions is smoothed with a Gaussian kernel to produce a more realistic probability density function. The implementation of such Gaussian smoothing improved the average mean absolute rotational error from 1.96° in the absence of smoothing to 0.48° with a smoothing kernel spanning 10 bins. However, the algorithm was robust to the width of the Gaussian kernel. Kernel widths of 6 to 16 bins produced similar registration precision ($\alpha = 0.05$). A kernel width of less than 6 bins deteriorated the precision and a width of more than 16 bins resulted in cases of unsuccessful registrations (mean absolute error $>$10°).

3.5 Conclusion

In this paper we proposed a novel method for 3D/2D registration. This method achieved the accurate registration through matching the probability density function of the gradient directions of the digitally reconstructed images and experimental images. This density function was estimated through a weighted histogram of gradient directions of the images, which was smoothed by a Gaussian kernel.

Our method is an improvement over the existing ones in that it is significantly more robust to the initial distance from the true registration and requires much less computational cost. The quality of the registration of our method also showed independence from the initial solution as long as the solution was within 90° of the true registration. In the experiments, our method was able to reach the registration accuracy of an established existing method (weighted edge matching score) using only 5% of the DRRs (10% of the computation time) and with a significantly larger distance from the true registration.

The proposed matching algorithm could be sensitive to the existence of irrelevant objects in the fixed image since these objects add substantial noise to the PDF of the gradient directions, potentially misleading the algorithm. It could be challenging to apply this algorithm for the registration of 3D bone models to fluoroscopic images with noise and additional tissues. We have resolved this problem through a pre-registration segmentation of the tissue of interest by performing a deformable 2D registration of an atlas DRR to the fluoroscopic image.

We have successfully used this method for the registration of 3D bone models to noisy clinical fluoroscopic images obtained from mini C-arm imaging machines. The results showed improvements over the existing methods as the registration was performed in reduced time and was less sensitive to the high amount of noise, which is usually associated with fluoroscopic images obtained from mini C-arms. Future work will involve a scientific reports of the performance of the method in such clinical settings.

Chapter 4

Fast and Accurate Joint Motion Analysis Using Mini C-Arm Fluoroscopy

4.1 Introduction

Reliable joint kinematic analysis is of crucial importance in many orthopedic applications related to surgery assessment, abnormality diagnosis, and prosthesis design. Such a method would aim at describing the motion of each joint bone by finding its 3D position at each time during the movement. The common in-vivo motion analysis method of using skin markers is not effective for orthopedic surgeries, because the soft tissue between the bone and the skin adds a considerable amount of distortion to the movement, making the results not precise enough for such purposes. The implantation of fiducials on the joint bone is also not practical for in-vivo applications due to the highly intrusive nature of the process and the distortion of the natural motion pattern. Therefore, the common method for joint kinematic analysis is the use of medical imaging technology for capturing the movement and calculating the position of the bones.

Due to the lack of a technology for capturing a 3D and dynamic image of the joint movements, the joint kinematic analysis is traditionally done using a static 3D image of the bone obtained from either computed tomography (CT) or magnetic resonance imaging (MRI) and a radiographic 2D video of the joint movement from fluoroscopy. In this method, in order to find the position of the bone for each frame of the video, 3D/2D image registration methods are employed to find the 3D position of the bone so that the 2D projection of the volume matches with the radiographic image. The assumption behind this method is that, if the 2D projection parameters are set up after the fluoroscopy imaging setting, the most similar 2D projection to 2D frame results from the closest pose of the volume to that of the real bone at the time the frame was captured. Therefore, a 3D/2D image registration of the bone volume will provide

the 3D motion pattern of the bone during the video.

The use of medical imaging techniques for joint kinematic analysis is, however, challenging and not widely accessible. This is not only due to the complicated nature of the method, but also the high cost of the required imaging machinery. A full-size C-arm fluoroscopy machine is highly expensive, which has made it practically impossible for most laboratories to conduct research on the subject of joint kinematic analysis. As a result, the research labs currently working on this topic are numbered.

A more affordable alternative for capturing fluoroscopy videos of joint movement is the mini C-arm fluoroscopy machine, which is a smaller version of the full-size one. Not only is this machine considerably less expensive than its full size version, but it is also more mobile and exposes the patient to a smaller dosage of radiation, leading to a lower probably of radiation side-effects. This category of imaging machines is however absent from the field due to lower image quality and the smaller field of view compared to the full size one. The additional amount of noise detracts from the quality of the image so much that the edges of the bones, which are one of the key elements for image registration, might not be discernible. In our experiments, the random nature of this noise showed it impossible to do useful noise reduction using white field correction [66]. Other noise reduction methods were also detrimental to the details of the image. Moreover, the smaller size of the image sensors in mini C-arms results in a much smaller field of view causing a smaller portion of the bones to be visible, which results in considerably less data to use for image matching.

In addition to the high cost, this method is time-intensive and user dependent. The time-intensity of the procedure stems from the fact that 3D/2D image registration involves the generation of many digital reconstructed radiographs (DRRs), which are complex to generate. The interaction with the user is required because most existing methods are able to perform registration only if the initial position of the bone is in the vicinity of the true registration and the algorithm will not converge otherwise. This requires the user to perform a visual registration of each bone prior to running the algorithm. Not only is this visual registration hard at times, the user can not know whether his registration is close enough for the algorithm to work prior to running it. This might necessitate multiple runs of the algorithm with different initial bone positions, which exacerbates the time intensity of the procedure.

In chapter 2 we introduced a DRR generation method, which solves the issue of large DRR generation time. In chapter 3, we introduced a new feature-based 3D/2D image registration algorithm. Not only was this method able to perform registration in a shorter time due to the significantly smaller number of generated DRRs, it was considerably less sensitive to initial position of the bone. In our experiments, this method was able to converge to registration from anywhere within 90 degrees of the true registration in terms of rotation. These properties make this method desirable for replacing the manual registration phase of the procedure. However, the feature used for image registration in this method, namely the weighted histogram of gradient directions, is highly sensitive to the existence of irrelevant object on the image. This impedes the method from being used in a real clinical settings as the radiographs not only involve other joint bones, but also soft tissue and image noise.

In this chapter we put forward a registration framework for 3D/2D image registration using the mini C-arm. Not only does this method make it possible to take advantage of the higher accessibility of the mini C-arm, it provides a method for isolation and reconstruction of the bone of interest through the deformable registration of an atlas to each frame. Having the bone isolated makes it possible to use the weighted histogram of gradient directions (WHGD) method, from chapter 3. This method could replace the time-consuming tinkering of the initial bone position mentioned above due to its robustness to the initial solution. And finally, an edge-based image registration method is applied to the results obtained from the WHGD, which improves the registration of the bone volume.

4.2 Method

4.2.1 Data Acquisition and Preprocessing

The 3D bone volumes in this study were obtained from CT image series of four human cadaver neck spines. In each neck spine, the C2 and C3 vertebrae were implanted with four or five fiducials to help calculate the ground truth registration. Then, CT scans of the neck spines were captured where the size of the pixels in the CT images was 0.256 mm by 0.256 mm with an inter-slice distance of 0.625 mm. On the 3D CT image, the pixels pertaining to the vertebrae of interest (C2 and C3) were labeled using ITK-Snap software [76]. Due to the

subtlety of the borders of the vertebrae, none of the available automatic segmentation methods resulted in acceptable precision and the segmentation was performed manually. The output of the segmentation process was two separate new volumes, one for C2 and another for C3. The pixels representing other tissues were set to zero resulting in volumes with a large number of background pixels. Since DRR generation time is a function of volume size, and background data does not contain any valuable information, the volumes were reduced in size by dropping the background pixels as follows.

All the vertebra volumes were rotated so that the 3D coordinate system of the image coincides with the principal axes of the vertebra. This coordinate system transformation serves several purposes:

- it provides some information, which could make a search in the rotation space easier. It is known that the scattering of a group of points - nonzero volume pixels - is maximum along the principal axis of the points with largest eigenvalue. This means that a rotation of the points around this axis results in the least amount of movement in the points compared to any other axis. Similarly, a rotation of the volume around its principal axis with smallest eigenvalue results in the largest change in the volume and therefore its projection. This could provide us with some guideline when searching the rotation parameter space by showing that a rotation around the principal axes with largest eigenvalue needs more effort because the effect of the same amount of rotation is more subtle compared to the other parameters.

- it helps the volume rotation and transformation be more visually meaningful, because this coordinate system is anatomically more meaningful.

- it helps with approximating the smallest bounding box around the vertebra and therefore eliminating more background pixels, which leads to increasing the speed of DRR generation.

- it unifies the orientation of the vertebrae across different subjects.

After this change of coordinate systems, the volume is shrunk to the smallest bounding box

around the vertebra. This results in a substantial reduction in volume size leading to considerable speed increase in DRR generation. In the resulting image, the fiducials are segmented in a similar method.

The motion of human neck spine was simulated and captured using a mini C-arm machine. For each spine, two flexion and two extension movements were manually simulated and recorded. A roughly lateral view was kept to across all videos. The videos had a minimum of 75 and a maximum of 120 frames. By experience, the first few frames of the videos have shown to suffer from such a high amount of noise and fluctuation in intensity that renders the frames useless. Therefore, in all of the captured fluoroscopy videos, the first 20 frames are cut out. Afterwards, each fluoroscopy video was corrected for the image distortion that is intrinsic to this system of imaging. Also, the location of the camera and the spacing of the image were extracted through camera calibration. The details of distortion correction and camera calibration are explained in the following subsection.

4.2.2 Distortion Correction and Camera Calibration

Distortion Correction

The radiographic images obtained from C-arm fluoroscopy machine are not an exact representation fo reality due to distortion. This distortion is an intrinsic feature of the image intensifiers used in fluoroscopy machines. Previous work has shown that this distortion could be substantial, especially near the periphery of the field of view [57, 62]. In order to use such images in applications where precision in spatial measurements of the image is important, these images need to be corrected for this distortion.

Fluoroscopic images suffer from two sorts of distortion: pincushion-, and S-shaped distortion. The most visible distortion in such images is the pincushion distortion, which is formed due to the mapping of a flat image onto a curved input phosphor [62]. This kind of distortion is intrinsic to the image intensifier and is independent of its positioning. The secondary form of distortion is an S-shape curve in the image, which is caused by the magnetic field of the earth and the objects around the intensifier. This distortion is dependent on the orientation of the image intensifier and could change if the C-arm is rotated or moved [44]. In this work, these

two different distortions are treated as one complex distortion [72].

The common method for the correction of image distortion is to estimate the distortion pattern with a mathematical function by using a phantom grid to learn the parameters of this function [15, 16]. Then, an inverse of the obtained deformation matrix could be applied to each fluoroscopy image to provide a realistic image. In this work, we used a bi-planar phantom grid for this purpose. The design map of this phantom was an adaptation of a similar design for the distortion correction of a full-size C-arm machine. The lower plane of the phantom grid consists of beads of two different sizes located at equal distances from each other. This plane is positioned exactly on top of the intensifier and the location of the beads provides the sample points required for the learning of distortion function parameters (Figure 4.1, left). The upper plane of the phantom includes lines, which are visible in the image. These lines provide the data necessary for the camera calibration (Figure 4.1, middle). Together, one calibration image with these two planes provides us the data which is required for both image distortion correction and camera calibration (Figure 4.1, middle).

In the first step, the beads and the lines need to be separated in the calibration image. To do so, first the image is smoothed to compensate for the noisy nature of the imaging system. Then, the connected components on the image are detected and labeled. The area of the components in terms of pixels provides us with a simple, but effective clue as to whether the component is a line or a bead.

To use the beads as training sampling points, the correspondence between the beads in the image and those of the design map has to be found. This is done by finding the affine registration of the design map to the image. First, large and small beads are separated using the area of each component representing a bead. Then the central bead in the image is detected as the large bead with minimum distance from the farthest large bead to it. The location of the central bead on the image and the design map provides the translation between the two images.

To find the rotation between the image and the design map, the following normal equation is solved for T:

$$T.C_d = C_i$$

where C_d is the location of the large beads in the design map and $C - i$ is their location on

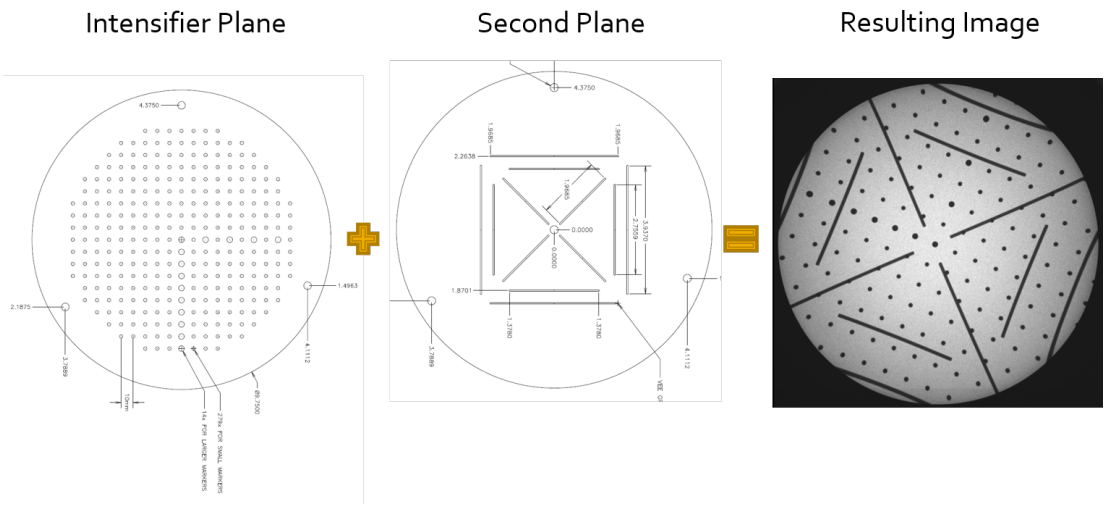


Figure 4.1: The distortion correction and camera calibration for the fluoroscopy images is done usuing a bi-planar phantom grid. The lower plane of the phantom contains fiducials located at equal distances from each other (left). Since this plane is positioned exactly on top of the image intensifier, no perspective distortion of the distances is involved and the deformation of the fiducial locations is caused by intensifier distortion only. The distortion pattern of the intensifier will learned using this plane. The second phantom plane is located at a distance above the first plane. The lines on this plane provide the information required for calculating the position of the rays source. Both distortion correction and camera calibration is performed using on image of the phantom, which includes data from both of the planes.

the image. Matrix T is the transformation matrix which matches the large beads on the design map to those on the image. This matrix, however, is not a purely rotation matrix. This is due to the scaling between the two images and also a small amount of error. Since, due to distortion, the scaling of the bead locations is not uniform in the image, we are merely interested in the rotation angle and not the scaling parameters. To separate the scaling from the rotation, an SVD decomposition of the matrix is calculated and the rotation matrix is computed by dropping the matrix of eigenvalues.

$$T \rightarrow S.V.D$$

$$\text{rotation matrix} = S.D$$

The scaling factor is, instead, calculated using the distance between the central bead and its closest neighbors. This is due to the fact that distortion is minimum at the center of the image.

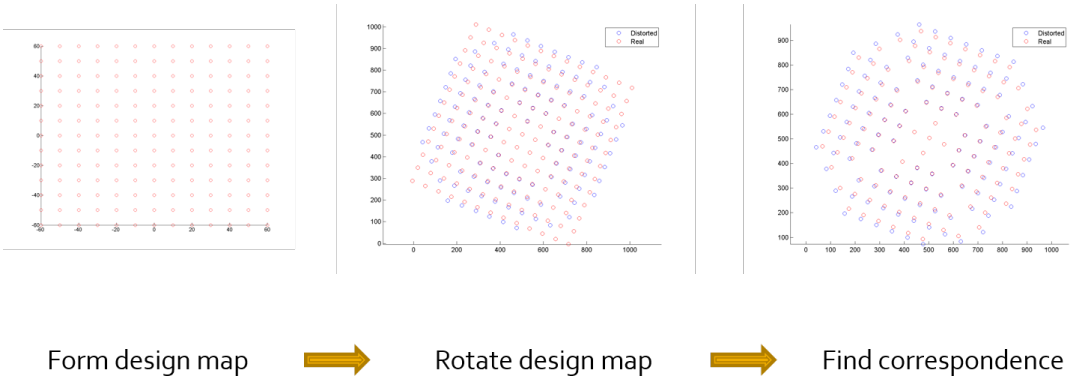


Figure 4.2: Left: the designed location of the beads on the lower plane of the phantom. Middle: the translation and rotation difference between the beads in the image and the designed map is calculated and the correspondence of the image and designed map beads is found. Right: the image beads covered by the calibration lines are detected and removed. The distortion of the location of the beads in the image compared to their ideal locations on the design map is visible. This difference is used to find the distortion pattern of the fluoroscopy images and revert it.

After the registration is finished and the transformation is applied to the design map, the correspondence between the beads in the design map and the image is found by finding the nearest design map bead to each bead on the image (Figure 4.2).

We are looking for the function which maps each point on the distorted fluoroscopy image to its real location on the image if there was no distortion. Now that the correspondence between the points on the image and their real location on the design map is available, we can use interpolation to estimate such a function. While this interpolation will give us the parameters of such a function, the type of function remains up to us to chose. The choice of function has been investigated in the literature of the field and it is concluded that a linear approach with Bezier polinomials results in satisfactory results [28]. In this work we used the same interpolation framework and and parameters as in [15] and [16]. In this framework the location of the beads in the design map (u_d, v_d) is seen as a function of their location on the image (u_0, v_0) and the distortion correction function is interpolated as

$$\Delta d = (\Delta u, \Delta v) = (u_d, v_d) - (u_0, v_0)$$

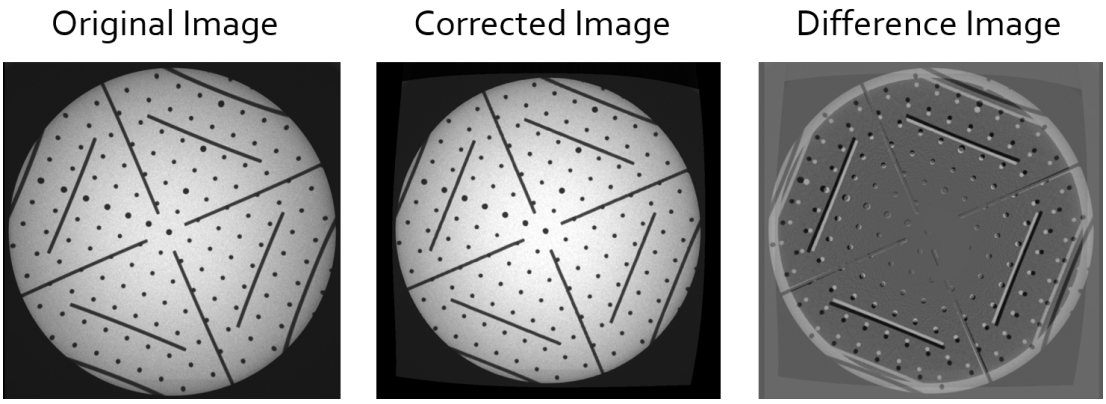


Figure 4.3: Left: the image of the phantom. The distortion pattern of the images is calculated using the bead locations on this image. The curvy nature of the distortion is visible in the slight bending of the calibration lines. Middle: the phantom image after distortion correction. The bending of the calibration lines is fixed. Right: the difference between the phantom images before and after distortion correction. This image shows the extent of distortion, which is especially more intense towards the edges of the sensor.

where $B_{ij}(u_0, v_0)$ are the Bezier polynomials:

$$(u_d, v_d) = \Sigma_{i=0}^{n} \Sigma_{j=0}^{n} C_{ij} B_{ij}(u_0, v_0)$$

. After the coefficients of the polynomials, C_ij, are found each new fluoroscopy image is corrected based on the interpolated function 4.3.

$$B_ij(u_0, v_0) = \binom{n}{i} u_0^i (1 - u_0)^{n-i} \binom{n}{i} v_0^j (1 - v_0)^{n-j}$$

Camera Calibration

After the calibration image is corrected for distortion, it could be used for camera calibration. The equation of the lines segmented in the distortion correction phase is extracted using registration and the intersection of the lines is calculated. The calculated coordinates of the line intersections are, in fact, the position of the projection of the line intersections on the image plane, which approximately coincides with the lower plane of the calibration phantom. The real location of the line intersections is known from the design map of the phantom. Each pair

of projected line intersection and the real location define a line that has to cross the ray source. Therefore, the intersection points result in a 3D system of equations, the solution of which is the location of the ray source.

4.2.3 Registration

Deformable Registration of the Atlas

In order to benefit from the advantages of the method introduced in chapter 3, the object of interest needs to be isolated from other tissue and background. This is however not a trivial problem, because the perspective projection of the neck onto the image plane causes depth information along the projection axis to be lost. A great amount of intensity and information is lost due to overlap with other tissues, low image quality, or noise. Therefore, the traditional methods of object segmentation for 2D images are not necessarily effective for the segmentation of the vertebrae. To achieve this, we used an atlas-based method, which not only helps isolate the vertebra from the rest of the image approximately, but also revives those edges of the bone that are lost due to overlap with other tissue.

The atlas is a fluoroscopy frame, for which the ground truth registration is known a priory. The true registration parameters for the atlas are found using the fiducials implanted in the bone. The fiducials on the atlas are detected using Hough transform for circle detection [67] or manually depending on the quality of the frame. Afterwards, the registration parameters are found so that the distance between the centers of the fiducials on the fluoroscopy and the 2D projection of their counterparts in the CT volume is minimized. The projection parameters are the same as used in DRR generation and the minimization is done through genetic [5] and downhill simplex algorithms [53, 47].

To locate the vertebra of interest the atlas was deformably registered to every frame of the fluoroscopy video of the subjects (Figure 4.4). The deformable registration was performed using the Advanced Normalization Tools [3] (http://picsl.upenn.edu/software/ants). The energy function for the registration process was chosen as mutual information [54] with 32-bin histograms. To shorten the registration time, a three-level multi-resolution coarse-to-fine approach was employed [22]. Since a total non-rigid registration of the atlas to each frame might

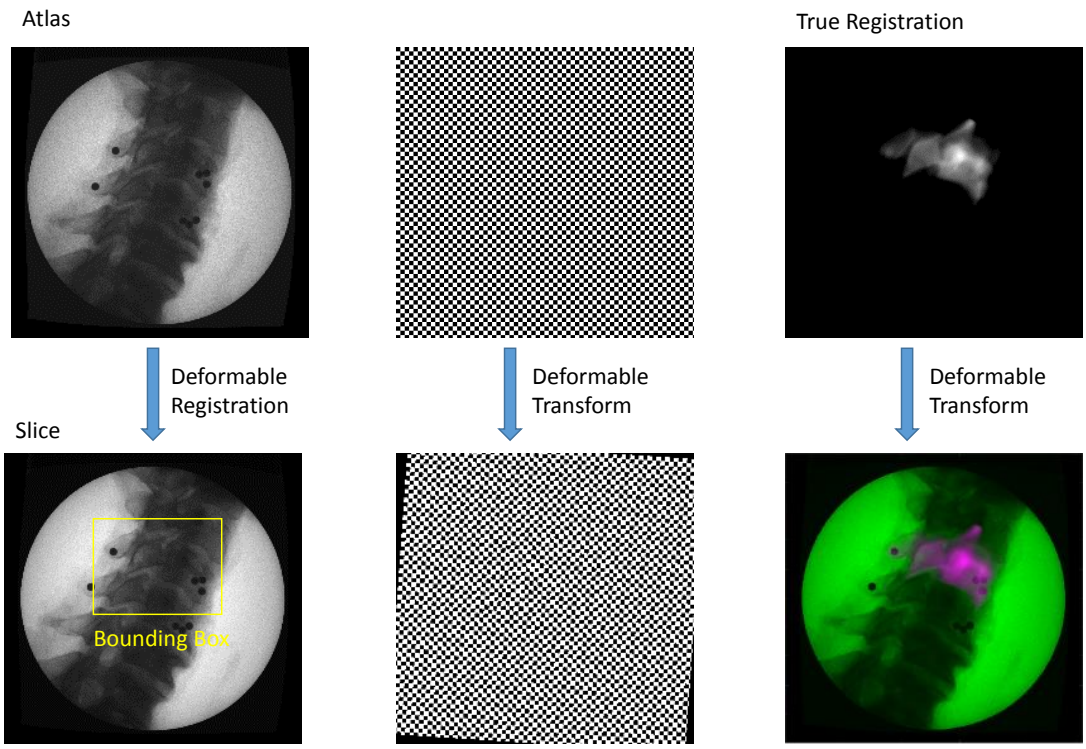


Figure 4.4: The bone of interest is isolated from the rest of the image using a deformable registration of an atlas to each frame. The transformation obtained by this registration is then applied to the DRR of the true registration of the atlas. The deformed DRR is then used as an approximation for the isolated bone. Through this method those edges of the bone which are lost due to low image quality and overlap with other tissue are also reconstructed.

be computationally intensive, a binary mask was used to limit the registration to a bounding box for the vertebra. The bounding box for every frame was calculated using the registered vertebra in the last frame assuming that the location of vertebra will not have an abrupt change in between frames.

3D/2D Registration

The 3D/2D registration of the bone volumes to the deformed atlas DRRs was performed as explained in chapter 3. The registration process consists of a sequence of optimizations for registration parameters. The optimization is mainly done on out-of-plane rotation parameters,

namely X and Y rotation. At each iteration, a new set of out-of-plane parameters is used to generate a DRR. The weighted histogram of the gradient directions is then formed and compared to that of the fixed image. The amount of shift in the histogram, which results in the highest similarity measure between the two histograms is then taken as the out-of-plane rotation parameter. Each time the best observed similarity measure since the start of the algorithm is updated, the translation parameters are also updated and used for the next generated DRRs. The space of X and Y rotation parameters is searched using a pattern search strategy (Chapter 3). Prior to the algorithm, the initial solution is improved according to the sampling strategy explained in the previous chapter.

The aforementioned 3D/2D registration method has properties, which make it highly favorable for replacing the manual initialization of the registration, however a more precise registration method is still needed for a fine registration of the volume. To have such a fine tuning step in our framework, we developed a similarity measure based on matching the edges in the fixed and the moving images. The idea of edge matching for registration is one of the most established registration ideas in the filed. While such methods are precise, they are myopic and need to be close to the true registration to work. Also, their performance depends on the clarity of the edges on both images, which is not the case in images obtained from the mini C-arm fluoroscopy machine. In our similarity measure, we tried to address these shortcomings by using a distance map and also incorporating the atlas DRRs into our registration process.

The similarity measure is calculated using the DRR, the fluoroscopy frame, and the deformed atlas DRR. All images are normalized and their edges are detected using the canny edge detection method [13]. Thereafter a distance transform of the edges on the deformed atlas DRR is calculated, where the intensity of pixels is equal to the Euclidean distance of that pixel from the closest non-zero pixel on the edge image [36]. To emphasize the intensity of the pixels closer to edges, all pixel intensities in the distance map are then squared. this also provides an image of non-negative pixel intensity values. The resulting image is then multiplied with the edge image of the fluoroscopy frame in a pixel-wise fashion. The edges on the fluoroscopy image are weighted according to their length to give less significance to smaller edges. From here on, each time a new DRR is generated, the similarity measure is calculated as the sum of pixel intensities on the image which is the result of the multiplication of the canny edges on the

DRR image and the image mentioned above.

The use of the distance map in the similarity measure is beneficial in two ways. First, the edges on the deformed atlas DRR serve as a weighting system, which puts more emphasis on those edges which are closer to the vetebra edges and might belong to it with a higher probability. This is especially useful because a good portion of edges are usually lost in the fluoroscopy image and a large amount of irrelevant edges exist, which don't belong to the vertebra. Second, the distance map addresses one of the problems with edge-based similarity measures, which is such measures are myopic. This class of measures are a good representation of how the two images are close to each other when the edges are close to matching each other. However if the images are far from one another, the overlap of their edges is a small number close to zero. In other words, such similarity measure do not give much information as to how far the images are from each other and therefore, they do not provide much guidance as to how to move towards registration when the moving and the fixed image are far from on another. The use of the distance map spreads the edge information to all pixel images. The intensity of each pixel shows its distance from the closest edge. This means that the similarity measure not only shows whether the two images match or not, it gives useful information as to how far the two images are from matching. This provides the possibility of moving towards registration even from far images by consistently maximizing the similarity measure. This maximization is done using the gradient-free downhill simplex optimization method [49].

4.3 Results

Registration was tested on a total of sixteen flexion and extension videos from C2 and C3 neck vertebrae for the four subjects. In the experiments, our method was able to reach a mean absolute error ($\pm$SD) of 0.63 ($\pm$0.25), 0.65 ($\pm$0.27), and 2.27 ($\pm$1.18) millimeters for X, Y, and Z translation and 2.42 ($\pm$1.19), 2.58 ($\pm$0.9), and 2.18 ($\pm$1.31) degrees for X, Y, and Z rotation parameters (Table 4.1). The mean ($\pm$SD) value of (signed) error was -0.62 ($\pm$0.28), -0.64 ($\pm$0.26), and -2.23 ($\pm$1.27) millimeters for X, Y, and Z translation and -2.37 ($\pm$1.27), -2.58 ($\pm$0.91), and 2.06 ($\pm$1.49) degrees for X, Y, and Z rotation.

The computation time for the registration of the volume to each fluoroscopy frame was an

average of 15.87 seconds. The registration time for the middle frames was considerably shorter than the first frame as the registration solution of each frame served as an initial solution for the next one. The average computation time for the deformable registration of the atlas to each fluoroscopy frame was 0.13 seconds with a standard deviation of 0.14 seconds. This short time was due to limiting the registration to the bounding box as explained in section 4.2. A total registration of the two images would result in considerably longer registration times and possibly worse results.

The use of the edge-based registration method for a fine tuning of the registration results showed to improve these results significantly. To verify this effect of this step, a set of registrations was run by skipping the edge-based registration and registering the 3D volume to the deformed atlas DRR. This registration was done on the C2 and C3 vertebrae from two subjects and the mean ($\pm$1.18) absolute error was 0.98 ($\pm$0.9), 1.39 ($\pm$1.12), and 2.41 ($\pm$1.98) millimeters for X, Y, and Z translation and 3.58 ($\pm$2.08), 3.94 ($\pm$1.56), and 3.68 ($\pm$2.12) degrees for X, Y, and Z rotation parameters. The results show the usefulness of applying the edge-based registration on the results. In our experiments, the inter-subject and inter-vertebra difference of the results did not show to be of statistical significance ($\alpha = 0.05$).

4.4 Discussion

In this chapter we provided a 3D/2D image registration framework for joint kinematic analysis in clinical settings. Joint kinematic analysis is an important medical procedure, which has many applications in orthopedic surgery. The kinematic analysis of joint movement is usually done through a 3D/2D registration of a static 3D bone volume obtained from CT image series to a 2D video of the joint movement captured by fluoroscopy. This procedure is however time consuming, user dependent, and expensive due to the high cost of a fluoroscopy machine. In this chapter we addressed these problems by suggesting a framework for performing the motion analysis using mini C-arms, which are the smaller and much less expensive versions of full size fluoroscopy machines.

To reduce registration time, the traditional method of DRR generation, namely ray casting, is replaced with the attenuation box method (chapter 2). In this method the attenuation value

of rays are precalculated by sampling the surface points of the volume and preprocessing the possible rays between every combination of two such points. At DRR generation time, the value of the rays is simply retrieved, which results in a major simplification of the processing and saving in time.

To further improve the speed and also reduce the dependence of the algorithm's performance on the initial solution set by the user, we used the feature based 3D/2D image registration described in chapter 3. We used the deformable registration of an atlas to each frame to isolate the bone of the interest from the rest of the image and also reconstruct the edges covered by other tissue. This isolation is necessary due to the fact that the existence of other tissues heavily misleads the weighted histogram of gradient directions, which is used as a feature for image registration in the method from chapter 3. Moreover, the images from mini C-arm machines are highly noisy and low quality which has kept this category of imaging machines from being used for joint kinematic analysis. The use of the atlas helps eliminate the noise and get close to high quality representation of the bone, which shares the same modality as the DRR which is going to be registered to it.

Since the reliance of the method to the deformed atlas might result in lower precision of the results, the registration results are improved by applying an edge-based registration method to the results. Such methods are usually effective when in the proximity of the true registration, but fail to converge when the initial solution is far from it. Since the registration is partly solved in the previous step, the use of such methods guarantees a higher quality of the results. In this method we used the distance map of the edges to provide the similarity measure with information on how far it is from the answer and smoothing the solution space. In our experiments, the use of this algorithm showed to improve the precision of the results without taking too much of solution time.

In our experiments, this method resulted in a mean ($\pm$SD) absolute error of 1.18 ($\pm$0.57) millimeters for translation and 2.39 ($\pm$1.13) degrees fro rotation parameters. The translational error along the Z-axis (out of plane translation) was considerably larger than the in-plane translational errors (X, and Y axes). Namely, this error was 2.27 millimeters compared to an average in-plane translational error of 0.64 millimeters. This result is in consistence with previous body of work on 3D/2D image registration [46, 64, 65], where the out-of-plane translational error

has proved to be the hardest translational parameter to register for. This is due to the fact that, while an in-plane translation has a clear effect on the DRR, the effect of an out-of-plane translation is subtle and only large translations of this kind manifest visible changes in the projected image. Although the rotational precision of the outcome is still not on par with results obtained from a full size fluoroscopy machine, it is still high enough for many orthopedic applications. Due to the substantially larger cost of a full size C-arm machine, this makes a mini C-arm a more favorable option for such applications, where an extreme precision of the results is not necessary.

The calculation time for registration in our experiments was low. The time for the deformable registration of the atlas to each frame was less than one second on average and the registration time of the volume was 15.87 seconds per frame. This high speed is thanks to two properties of our method. First, the DRR generation method described in chapter 2 results in an elimination of the biggest portion of computation time in 3D/2D registration algorithms. Second, the use of the registration method from chapter 3 leads into a much faster registration speed as described in the chapter. Also, the deformable registration of the atlas is expedited by using a bounding box for registration, which results in major saving of time.

Table 4.1: Average registration results for each vertebra under study. The eight vertebrae belong to four human cadaver subjects and for each vertebrae two flexion and two extension fluoroscopic videos were captured.

		Rotation (°)			Translation (mm)		
		X	Y	Z	X	Y	Z
Subject1-C2	MAE	+2.382	+2.572	+2.214	+0.602	+0.624	+2.612
	SDE	+1.402	+1.145	+1.061	+0.190	+0.200	+1.445
	ME	-2.321	-2.572	+2.143	-0.602	-0.624	-2.539
Subject1-C3	MAE	+2.463	+2.544	+2.693	+0.639	+0.621	+2.422
	SDE	+1.472	+0.435	+1.661	+0.151	+0.113	+1.354
	ME	-2.411	-2.544	+2.570	-0.639	-0.621	-2.391
Subject2-C2	MAE	+2.219	+2.513	+2.385	+0.656	+0.706	+2.076
	SDE	+1.399	+0.783	+1.455	+0.595	+0.448	+1.339
	ME	-2.167	-2.513	+2.324	-0.556	-0.689	-1.997
Subject2-C3	MAE	+2.469	+2.428	+1.920	+0.600	+0.624	+2.083
	SDE	+1.304	+1.463	+1.175	+0.074	+0.158	+0.905
	ME	-2.451	-2.386	+1.892	-0.600	-0.624	-2.060
Subject3-C2	MAE	+2.370	+2.527	+2.342	+0.653	+0.660	+2.079
	SDE	+1.083	+0.747	+1.861	+0.120	+0.255	+1.491
	ME	-2.366	-2.527	+2.165	-0.653	-0.660	-1.979
Subject3-C3	MAE	+2.797	+2.901	+2.194	+0.657	+0.646	+2.365
	SDE	+1.826	+0.933	+1.622	+0.330	+0.156	+0.817
	ME	-2.658	-2.901	+1.987	-0.657	-0.646	-2.365
Subject4-C2	MAE	+2.309	+2.621	+2.088	+0.628	+0.676	+2.385
	SDE	+0.759	+0.860	+1.612	+0.307	+0.438	+1.246
	ME	-2.309	-2.621	+1.859	-0.623	-0.663	-2.360
Subject4-C3	MAE	+2.347	+2.603	+1.575	+0.629	+0.601	+2.172
	SDE	+0.419	+0.482	+1.194	+0.147	+0.238	+1.371
	ME	-2.347	-2.603	+1.523	-0.629	-0.601	-2.136

Chapter 5

Conclusion and Future Research

Joint motion analysis is essential in orthopedics surgery. This is because the goal of most orthopedic surgeries is to restore the natural movement of the bones and it is not possible to measure the success of the surgery without the possibility of describing the joint movement. The common methods for kinematic analysis, such as skin mounted markers or bone implanted fiducials are not effective in orthopedics due to either the low precision of the results because of soft tissue or highly invasive nature of the method for in-vivo analysis. This necessitates the use of medical imaging technology for describing the 3D motion of the joint bones.

This class of kinematic analysis is traditionally performed using images obtained from CT scans and fluoroscopy images. To find the six degrees of freedom of each bone in the joint, a 3D model of the bone, which is reconstructed from CT image series of the joint, is matched to each frame of a fluoroscopy video of the joint motion captured in clinical settings. This image matching is performed using 3D/2D image registration methods, which try to find the 3D position of the bone volume, which results in the most similar 2D projection of the bone to the clinical 2D image. The assumption behind this method is that, if the parameters of the 2D projection are set after the camera parameters for in the clinical setting, the most similar 2D projection to the clinical video frame should result from the most similar position of the bone volume to that of the real bone at the time when the frame was captured. Finding such a position will provide the six degrees of freedom for the whole movement.

The registration process mentioned above is however slow, user dependent, and costly. The registration of the volume to each frame is a time-consuming process. This is due to the fact that each 2D projection of the volume, namely digitally reconstructed radiograph (DRR), is complex process requiring many memory access operations. The registration of the volume to each frame requires the generation of hundreds of DRRs, which lead to the time intensity

of the registration process. Also, current registration processes are highly user-dependent as they converge towards the true registration only when starting from a small distance from it. This necessitates the initialization of the algorithm by the user, which is problematic not only because a visual registration of the bone is tricky at times, but also because it is not possible to know whether the initial solution is close enough until a whole run of registration is completed. Another challenging aspect of this method is that a full size fluoroscopy machine is so expensive that few medical laboratories can afford to conduct analyses of this category.

In this thesis, we address the challenges mentioned above. We used a mini C-arm, which is a smaller and considerably more affordable version of the full size C-arm fluoroscopy machine. Although the mini version is cheaper and more mobile than the fluoroscopy machine and also exposes the patient to a less amount of radiation, its use in this context has been rare due to the lower image quality and field of view. In this thesis, we developed a registration framework, which is able to perform registration despite these factors. We also expedited the registration process through using a novel and fast DRR generation method in addition to a new feature-based registration method, which requires a substantially less number of DRRs. This registration is highly robust to the initial position of the bone, which minimizes the dependence of the procedure on the user.

In chapter 2 of this thesis, we introduced a new method for the fast generation of digitally reconstructed radiographs (DRR) from the 3D bone models. DRRs are synthesized radiographic images, which are produced by simulating the fluoroscopy imaging process. Since 3D/2D registration is performed using iterative optimization methods, a large number of DRRs need to be generated to search the solution space. This causes the registration process to be impractically lengthy, which detracts from the attractiveness of this method for clinical applications. In our method, we sampled the attenuation value of rays going through the volume by identifying them with their two intersection points with the surface of the volume. This method has the advantage of supporting any rigid transformation of the volume, which is not the case in existing methods of this category. Moreover, the use of memory in this method is efficient as no memory is assigned to rays which do not cross the volume.

Our DRR generation method showed to be able to produce high quality DRRs in significantly reduced time. In our experiments, we reduced DRR generation time by up to two order

of magnitude through preprocessing of the ray values. The reduction in DRR generation times becomes even more significant with an increase in the size of the volume, however this will come at the cost of memory and preprocessing time. In our experiments, the memory and preprocessing requirements were small and and the method will be effective for 3D models of subjects that are small enough to fit the mini, or full size, C-arms. An investigation of the methods effectiveness on larger volumes could be a subject of future study. The quality of the DRR was a function of the ray sampling resolution. A higher DRR quality will not necessitate larger DRR generation time rather larger memory and preprocessing. We used the method in this chapter for DRR generation all through the experiments in this thesis.

In chapter 3, a novel feature-based method is introduced for general 3D/2D image registration. This method uses the probability density function of the image gradient directions as an image feature to register the images with a substantially reduced number of necessary DRRs. The properties of this feature allow for the six degrees of freedom to be optimized consequently rather than simultaneously. This results in a major reduction in time, because the complexity of searching a six dimensional solution space is far higher than that of a sequence of one or two dimensional ones. In addition to increased speed, this method is highly insensitive to the initial position of the bone. This is due to the fact that the reduction in the dimensions of the search space makes it feasible to do a sampling of the solutions. Using this method, we were able to register the 3D volume to the 2D image with initial solutions as far as 90°from the final position, which is a substantial improvement over similar existing methods.

In chapter 4, we developed a registration framework to take advantage of the methods developed in previous chapters in a clinical setting. To be able to use the feature explained in chapter 3 for registration, the bone of interest was segmented from the rest of the image using a deformable registration of an atlas to each frame. This process not only eliminates the distortion of the feature caused by other tissue and image noise, it helps with reviving bone edges which are lost due to the low quality of the image or overlap with other tissues. This deformation however leads in lower precision in registration results, which is addressed by applying an intensity-based registration of the results. In the second registration, a distance map of the canny edges in the fixed image is used to form a similarity measure resulting in a smooth and consistent move towards optimization. Edge-based registration methods are highly effective

when in the vicinity of the true registration, but have a low convergence span. In our method, we used the feature-based registration phase as a guarantee that the initial solution to the second registration is close to the true answer.

In our experiments, the method was able to reach the true registration in short time. The mean ($\pm$SD) absolute registration error was 1.18 ($\pm$0.57) millimeters and and 2.39 ($\pm$1.13) degrees fro translation and rotation parameters respectively. Although higher precisions are reported in the literature by using the full size C-arm fluoroscopy machine, these results are highly attractive as they are good enough for many medical applications and they will allow a much broader set of laboratories to have access to such analysis. The speed of registration was high as registration for each frame took and average of less than 20 seconds. This speed could be even more enhanced if the programming is optimized as our experiments were done mostly in Matlab (except for DRR generation).

Future research could aim at a more integrated framework for the whole process flow. Although the registration part of the kinematic analysis is mostly automatic, there is a large body of preprocessing, such as segmentation and calibration, that needs to be implemented before the registration. Future work should focus on automatizing of the whole process. An investigation of the methods effectiveness on other body joints is also a promising direction. The 3D/2D registration is intensively challenging for more delicate joints such as human wrist or for bones with a symmetric structure such as the elbow radial bone. Also, it is expected that in a more commercial setting with large numbers of subjects, a multi-atlas approach will be more viable as the variety of the subject is much higher. A GPU (graphic processing unit) based implementation of the method could also contribute to an even faster implementation as it is extensively used in similar applications.

References

[1] Stacey Acker, Rebecca Li, Heather Murray, Paul St John, Scott Banks, Shang Mu, Urs Wyss, and Kevin Deluzio. Accuracy of single-plane fluoroscopy in determining relative position and orientation of total knee replacement components. *Journal of biomechanics*, 44(4):784–787, 2011.

[2] Ethem Alpaydin. *Introduction to Machine Learning*. The MIT Press, 2nd edition, 2010.

[3] Brian B Avants, Charles L Epstein, Murray Grossman, and James C Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.

[4] Mousumi Banerjee, Michelle Capozzoli, Laura McSweeney, and Debajyoti Sinha. Beyond kappa: A review of interrater agreement measures. *Canadian journal of statistics*, 27(1):3–23, 1999.

[5] Wolfgang Banzhaf, Peter Nordin, Robert E Keller, and Frank D Francone. *Genetic programming: an introduction*. Morgan Kaufmann Publishers San Francisco, 1998.

[6] Joseph Stiles Beggs. *Kinematics*. CRC Press, 1983.

[7] Marcel Berger and Bernard Gostiaux. *Differential Geometry: Manifolds, Curves, and Surfaces: Manifolds, Curves, and Surfaces*, volume 115. Springer Science & Business Media, 2012.

[8] Paolo Bifulco, Mario Cesarelli, Robert Allen, Maria Romano, Antonio Fratini, and Giulio Pasquariello. 2d-3d registration of ct vertebra volume to fluoroscopy projection: a calibrationmodel assessment. *EURASIP Journal on Advances in Signal Processing*, 2010:2, 2010.

[9] J Bijhold. Three-dimensional verification of patient placement during radiotherapy using portal images. *Medical physics*, 20(2):347–356, 1993.

[10] W. Birkfellner, R. Seemann, M. Figl, J. Hummel, C. Ede, P. Homolka, X. Yang, P. Niederer, and H. Bergmann. Wobbled splatting–a fast perspective volume rendering method for simulation of x-ray images from CT. *Phys Med Biol*, 50(9):73–84, May 2005.

[11] Marcel Breeuwer, John P. Wadley, HLT De Bliek, Johannes Buurman, PAC Desmedt, P Gieles, Frans A. Gerritsen, NL Dorward, ND Kitchen, B Velani, et al. The easi project-improving the effectiveness and quality of image-guided surgery. *Information Technology in Biomedicine, IEEE Transactions on*, 2(3):156–168, 1998.

[12] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376, 1992.

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

[14] Sung-Hyuk Cha. Comprehensive survey on distance/similarity measures between probability density functions, 2007.

[15] Gouthami Chintalapani, Ameet K Jain, and Russell H Taylor. Statistical characterization of c-arm distortion with application to intra-operative distortion correction. In *Medical Imaging*, pages 65092Y–65092Y. International Society for Optics and Photonics, 2007.

[16] Gouthami Chintalapani and Russell H Taylor. C-arm distortion correction using patient ct as a fiducial. In *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1180–1183. IEEE, 2007.

[17] William C. Davidon. Variable metric method for minimization. *SIAM Journal on Optimization*, 1(1):1–17, feb 1991.

[18] Douglas A Dennis, Mohamed R Mahfouz, Richard D Komistek, and William Hoff. In vivo determination of normal and anterior cruciate ligament-deficient knee kinematics. *Journal of biomechanics*, 38(2):241–253, 2005.

[19] Osama M Dorgham, Stephen D Laycock, and Mark H Fisher. Gpu accelerated generation of digitally reconstructed radiographs for 2-d/3-d image registration. *Biomedical Engineering, IEEE Transactions on*, 59(9):2594–2603, 2012.

[20] Jingfan Fan, Jian Yang, Feng Lu, Danni Ai, Yitian Zhao, and Yongtian Wang. 3-points convex hull matching (3pchm) for fast and robust point set registration. *Neurocomputing*, 194:227–240, 2016.

[21] Benjamin J Fregly, Haseeb A Rahman, and Scott A Banks. Theoretical accuracy of model-based shape matching for measuring natural knee kinematics with single-plane fluoroscopy. *Journal of biomechanical engineering*, 127(4):692–699, 2005.

[22] Soheil Ghafurian, Antong Chen, Catherine Hines, Belma Dogdas, Ashleigh Bone, Kenneth Lodge, Stacey O'Malley, Christopher T Winkelmann, Ansuman Bagchi, Laura S Lubbers, et al. Automatic pose correction for image-guided nonhuman primate brain surgery planning. In *SPIE Medical Imaging*, pages 97860O–97860O. International Society for Optics and Photonics, 2016.

[23] Soheil Ghafurian, Ilker Hacihaliloglu, Dimitris N. Metaxas, Virak Tan, and Kang Li. 3d/2d image registration using weighted histogram of gradient directions. In Robert J. Webster and Ziv R. Yaniv, editors, *Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*. SPIE-Intl Soc Optical Eng, mar 2015.

[24] Soheil Ghafurian, Dimitris N Metaxas, Virak Tan, and Kang Li. Fast generation of digitally reconstructed radiograph through an efficient preprocessing of ray attenuation values. In *SPIE Medical Imaging*. International Society for Optics and Photonics, 2016.

[25] KGA Gilhuijs, PJH Van De Ven, and M Van Herk. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Medical physics*, 23(3):389–399, 1996.

[26] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[27] A Ardeshir Goshtasby. *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*. John Wiley & Sons, 2005.

[28] AP Hammersley, SO Svensson, and A Thompson. Calibration and correction of spatial distortions in 2d detector systems. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 346(1):312–321, 1994.

[29] Robert Hooke and T. A. Jeeves. " direct search" solution of numerical and statistical problems. *J. ACM*, 8(2):212–229, April 1961.

[30] Sheng-Chieh Huang, Liang-Gee Chen, and Hao-Chieh Chang. A novel image compression algorithm by using log-exp transform. In *Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium on*, volume 4, pages 17–20 vol.4, Jul 1999.

[31] Harold Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 186(1007):453–461, 1946.

[32] Hans J. Johnson, M. McCormick, L. Ibáñez, and The Insight Software Consortium. *The ITK Software Guide*. Kitware, Inc., third edition, 2013. *In press*.

[33] Avinash C.. Kak and Malcolm Slaney. *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.

[34] P Kazanzides, B Williamson, RH Taylor, et al. Anatomy-based registration of ct-scan and intraoperative x-ray images for guiding a surgical robot. *Medical Imaging, IEEE Transactions on*, 17(5):715–728, 1998.

[35] Erwan Kerrien, M-O Berger, Eric Maurincomme, Laurent Launay, Régis Vaillant, and Luc Picard. Fully automatic 3d/2d subtracted angiography registration. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI99*, pages 664–671. Springer, 1999.

[36] Ron Kimmel, Nahum Kiryati, and Alfred M Bruckstein. Sub-pixel distance maps and weighted distance transforms. *Journal of Mathematical Imaging and Vision*, 6(2-3):223–233, 1996.

[37] Yasuyo Kita, Dale L Wilson, and J Alison Noble. Real-time registration of 3d cerebral vessels to x-ray angiograms. In *Medical Image Computing and Computer-Assisted InterventationMICCAI98*, pages 1125–1133. Springer, 1998.

[38] Richard D Komistek, Douglas A Dennis, and Mohamed Mahfouz. In vivo fluoroscopic analysis of the normal human knee. *Clinical orthopaedics and related research*, 410:69–81, 2003.

[39] S. Kullback and R. A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22(1):79–86, 03 1951.

[40] STEPHANE LAVALLEE and Jocelyne Troccaz. 3 2 computer-assisted spinal surgery using anatomy-based. *Computer-integrated Surgery: technology and clinical applications*, page 425, 1996.

[41] L Lemieux, R Jagoe, DR Fish, ND Kitchen, and DGT Thomas. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Medical physics*, 21(11):1749–1760, 1994.

[42] Cheng-Chung Lin, Shuo Zhang, Jens Frahm, Tung-Wu Lu, Chao-Yu Hsu, and Ting-Fang Shih. A slice-to-volume registration method based on real-time magnetic resonance imaging for measuring three-dimensional kinematics of the knee. *Medical physics*, 40(10):102302, 2013.

[43] Yang Liu, Zhan-Li Sun, Ya-Ping Wang, and Li Shang. An eigen decomposition based rank parameter selection approach for the nrsfm algorithm. *Neurocomputing*, 198:109–113, 2016.

[44] Harel Livyatan, Ziv Yaniv, and Leo Joskowicz. Robust automatic c-arm calibration for fluoroscopy-based navigation: a practical approach. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 60–68. Springer, 2002.

[45] Mohamed R Mahfouz, Steven M Traina, Richard D Komistek, and Douglas A Dennis. In vivo determination of knee kinematics in patients with a hamstring or patellar tendon acl graft. *The journal of knee surgery*, 16(4):197–202, 2003.

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

[47] Ken IM McKinnon. Convergence of the nelder–mead simplex method to a nonstationary point. *SIAM Journal on Optimization*, 9(1):148–158, 1998.

[48] Martin J Murphy. An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery. *Medical Physics*, 24(6):857–866, 1997.

[49] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.

[50] Graeme P Penney, Philipp G Batchelor, Derek LG Hill, David J Hawkes, and Juergen Weese. Validation of a two-to three-dimensional registration algorithm for aligning pre-operative ct images and intraoperative fluoroscopy images. *Medical physics*, 28(6):1024–1032, 2001.

[51] Mark R Pickering, Jennie M Scarvell, and Paul N Smith. An improved ct to fluoroscopy registration algorithm for the kinematic analysis of knee joints. In *Digital Signal Processing, 2009 16th International Conference on*, pages 1–6. IEEE, 2009.

[52] Josien PW Pluim, JB Antoine Maintz, Max Viergever, et al. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

[53] Michael JD Powell. On search directions for minimization algorithms. *Mathematical programming*, 4(1):193–201, 1973.

[54] William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[55] Luis Miguel Rios and Nikolaos V Sahinidis. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3):1247–1293, 2013.

[56] T. Rohlfing, D. B. Russakoff, J. Denzler, K. Mori, and C. R. Maurer. Progressive attenuation fields: fast 2D-3D image registration without precomputation. *Med Phys*, 32(9):2870–2880, Sep 2005.

[57] Stephen Rudin, Daniel R Bednarek, and Roland Wong. Accurate characterization of image intensifier distortion. *Medical physics*, 18(6):1145–1151, 1991.

[58] D. B. Russakoff, T. Rohlfing, K. Mori, D. Rueckert, A. Ho, J. R. Adler, and C. R. Maurer. Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration. *IEEE Trans Med Imaging*, 24(11):1441–1454, Nov 2005.

[59] Daniel B Russakoff, Torsten Rohlfing, John R Adler, and Calvin R Maurer. Intensity-based 2d-3d spine image registration incorporating a single fiducial marker 1. *Academic radiology*, 12(1):37–50, 2005.

[60] J.N. Sarvaiya, S. Patnaik, and S. Bombaywala. Image registration using log-polar transform and phase correlation. In *TENCON 2009 - 2009 IEEE Region 10 Conference*, pages 1–5, Jan 2009.

[61] Dilip K. Sharma and D. S. Hooda. *Some Generalized Information Measures: Their Characterization and Applications*. LAP LAMBERT Academic Publishing, August 2010.

[62] Delia Soimu, Cristian Badea, and Nicolas Pallikarakis. A novel approach for distortion correction for x-ray image intensifiers. *Computerized Medical Imaging and Graphics*, 27(1):79–85, 2003.

[63] Derek F Stubbs. Three applications of neurocomputing in biomedical research. *Neurocomputing*, 2(2):61–66, 1990.

[64] Luca Tersi, Silvia Fantozzi, and Rita Stagni. 3d elbow kinematics with monoplanar fluoroscopy: in silico evaluation. *EURASIP Journal on Advances in Signal Processing*, 2010:14, 2010.

[65] Tsung-Yuan Tsai, Tung-Wu Lu, Chung-Ming Chen, Mei-Ying Kuo, and Horng-Chaung Hsu. A volumetric model-based 2d to 3d registration method for measuring kinematics of natural knees with single-plane fluoroscopy. *Medical physics*, 37(3):1273–1284, 2010.

[66] Vincent Van Nieuwenhove, Jan De Beenhouwer, Francesco De Carlo, Lucia Mancini, Federica Marone, and Jan Sijbers. Dynamic intensity normalization using eigen flat fields in x-ray imaging. *Optics express*, 23(21):27975–27989, 2015.

[67] Hough Paul VC. Method and means for recognizing complex patterns, December 18 1962. US Patent 3,069,654.

[68] Fei Wang and Baba C Vemuri. Non-rigid multi-modal image registration using cross-cumulative residual entropy. *International journal of computer vision*, 74(2):201–215, 2007.

[69] Jurgen Weese, Graeme P Penney, Paul Desmedt, Thorsten M Buzug, Derek LG Hill, and David J Hawkes. Voxel-based 2-d/3-d registration of fluoroscopy images and ct scans for image-guided surgery. *Information Technology in Biomedicine, IEEE Transactions on*, 1(4):284–293, 1997.

[70] Lee Westover. Footprint evaluation for volume rendering. In *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '90, pages 367–376, New York, NY, USA, 1990. ACM.

[71] E. T. Whittaker. *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies (Cambridge Mathematical Library)*. Cambridge University Press, 1989.

[72] Ziv Yaniv, Leo Joskowicz, Ariel Simkin, Maria Garza-Jinich, and Charles Milgrom. Fluoroscopic image processing for computer-aided orthopaedic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 325–334. Springer, 1998.

[73] Shihui Ying, Yuanwei Wang, Zhijie Wen, and Yuping Lin. Nonlinear 2d shape registration via thin-plate spline and lie group representation. *Neurocomputing*, 195:129–136, 2016.

[74] Byoung-moon You, Pepe Siy, William Anderst, and Scott Tashman. In vivo measurement of 3-d skeletal kinematics from sequences of biplane radiographs: application to knee kinematics. *Medical Imaging, IEEE Transactions on*, 20(6):514–525, 2001.

[75] Xinghu Yu, Ming Liu, Lingzhi Meng, and Liangbi Xiang. Classifying cervical spondylosis based on x-ray quantitative diagnosis. *Neurocomputing*, 165:222–227, 2015.

[76] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.