DOCTOR OF PHILOSOPHY

Registration of 3D models to 2D X-ray images using fast X-ray simulation and global optimisation algorithms

Wen, Tianci

Award date:
2023

Awarding institution:
Bangor University

PRIFYSGOL
BANGOR
UNIVERSITY

School of Computer Science and Electronic Engineering
College of Environmental Sciences and Engineering

Registration of 3D models to 2D X-ray images using fast X-ray simulation and global optimisation algorithms

---

Tianci Wen

Submitted in partial satisfaction of the requirements for the
Degree of Doctor of Philosophy
in Computer Science

Supervisor Dr. Franck P. Vidal

July 2023

Statement of Originality

The work presented in this thesis/dissertation is entirely from the studies of the individual student, except where otherwise stated. Where derivations are presented and the origin of the work is either wholly or in part from other sources, then full reference is given to the original author. This work has not been presented previously for any degree, nor is it at present under consideration by any other degree awarding body.

Student:

Tianci Wen

Statement of Availability

I hereby acknowledge the availability of any part of this thesis/dissertation for viewing, photocopying or incorporation into future studies, providing that full reference is given to the origins of any information contained herein. I further give permission for a copy of this work to be deposited with the Bangor University Institutional Digital Repository, the British Library ETHOS system, and/or in any other repository authorised for use by Bangor University and where necessary have gained the required permissions for the use of third party material. I acknowledge that Bangor University may make the title and a summary of this thesis/dissertation freely available.

Student:

Tianci Wen

Acknowledgements

I would like to express my deepest appreciation to all the people who made this thesis possible.

I would also like to extend my deepest gratitude to my research supervisor, Dr. Franck Vidal for providing unwavering guidance and insightful suggestions over time of researches and for his correction of this thesis.

I'm extremely grateful to my lovely wife, Ang Li for her unparalleled support and patience. The completion of this thesis would not have been possible without her.

Special thanks to Dr. Radu Mihail from Valdosta State University for providing essential research data and sharing related medical information.

Many thanks to all research radiographers including Dr. Jenna Tugwell-Allsup from Betsi Cadwaladr University Health Board, Graham Kelly from the Shrewsbury and Telford Hospital NHS Trust and Dr. Andrew England from University College Cork for conducting the survey of hand irregularities and for providing related medical information.

Abstract

Radiographs of the hand are useful in diagnosing and staging diseases such as rheumatoid arthritis (RA) and other musculoskeletal diseases. Radiographs are projections of the 3D anatomy, with the useful information such as pose and pathology becoming lost in the process. Recovering the 3D pose from a single radiograph allows detailed anatomical analysis in the 3D space and possible reduction of radiation exposure for the patients. The research area is around the registration of 3D mesh model to 2D X-ray images for the medical application of automatic disease diagnosis and treatment planning. The gap in knowledge was that many researches concentrate on registrations of 3D CT volume to radiographs whereas effect methods that work on 3D models are limited.

The aim of this thesis is to develop a 3D pose recovery framework for radiographs using a novel hybrid 3D/2D registration method. Our pose recovery pipeline consists of aligning a simulated digitally reconstructed radiograph of a 3D mesh model to a real radiograph. Our method heavily relies on fast X-ray simulations and optimisation algorithms. The method used to investigate the feasibility and performance of our approach was through conducting experiments of registrations using synthetic and real clinical data. We evaluated our approach using publicly available datasets including 15 radiographs of hands from the MURA dataset and a radiograph of the hip from the VHP, the Visible Man dataset. Results demonstrated that our approach works well in both registrations of radiographs with two different anatomical structure.

Our key findings were the registration of 3D mesh model to a single radiograph can be performed accurately and fast using our registration framework. Once suitable optimisation algorithm and metrics are identified, our approach tends to be reliable and registration results are similar

when repeated for several times. Also, our registration method can be easily modified to solve a specific registration problem where registration of a 3D polygon model to a single 2D image needs to be performed. The significance of these findings was that it clears the path for further development of a fully automatic process of clinical diagnoses and treatment planning of hand diseases. Further inquiry is required to evaluate the applicability and the performance of our registration approach to other articulated musculoskeletal anatomy, which would lead to the development of clinical applications for diagnosing and treatment planning of any disease targeting a particular anatomical region.

Contents

List of Figures

List of Tables

Acronyms

CBCT	cone beam computed tomography
CERN	European Organization for Nuclear Research
CG	conjugate gradient
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
CNN	convolutional neural network
CPU	central processing unit
CR	correlation ratio
CS	chi-square
CT	computed tomography
CTC	curve-to-curve
DICOM	Digital Imaging and Communications in Medicine
DL	deep learning
DRR	digitally reconstructed radiograph
DSA	digital subtracted angiograph
EA	evolutionary algorithm
EOD	entropy of difference
ES	Evolution Strategies
GC	gradient correlation
GD	gradient difference

GLSL	OpenGL Shading Language
GPU	graphics processing unit
IGE	image-guided endoscopy
IGRS	image-guided radiosurgery
IGRT	image-guided radiation therapy
MAE	mean absolute error
MAPE	mean absolute percentage error
MC	Monte Carlo
MI	mutual information
micro-CT	X-ray microtomography
ML	machine learning
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MSE	mean square error
NCC	normalised cross correlation
NIST	National Institute of Standards and Technology
NSGA-III	Non-dominated Sorting Genetic Algorithm III
NSGA-II	Non-dominated Sorting Genetic Algorithm II
PA	postero-anterior
PDM	point distribution model
PF	pareto front
PI	pattern intensity
PNG	Portable Network Graphics
PRS	pure random search

PTP	point-to-point
RA	rheumatoid arthritis
RC	rank correlation
RMSE	root mean square error
ROI	region of interest
SA	simulated annealing
SDD	source-detector distance
SLNC	sum of local normalised correlation
SNR	signal-to-noise ratio
SOD	source-object distance
SRC	stochastic rank correlation
SSD	sum of squared distance
SSE	sum of squared error
SSIM	structural similarity index measure
STC	surface-to-curve
STDEV	standard deviation
XMR	X-ray and MR interventional suite
ZNCC	zero-mean normalised cross correlation

Chapter 1

Introduction

1.1 Context

Radiographic (X-ray) views of hands help clinicians diagnose a wide range of conditions from simple fractures to skeletal maturity, and bone and joint damage due to autoimmune disorders. Simple radiograph imaging of the hand and wrist are often preferred over volumetric imaging modalities, such as computed tomography (CT) and Magnetic Resonance Imaging (MRI), due to lower patient radiation doses and the wide availability of equipment. Hand radiographs are 2D projections of 3D anatomy, therefore one view is often not enough for clinical diagnoses, so radiologists rely heavily on experience and atlas matching of multiple views to infer often small, but relevant irregularities. Alternatively, it can be achieved using 3D/2D registration methods (e.g. [175]) by registering the patient-specific model of hands to radiographs. Furthermore, a series of patient-specific models can be obtained using hand radiographs at different stages of the disease. We can monitor and model the changes from time to time for a specific hand disease in the 3D space. Thus, a fully automatic process of clinical diagnoses and treatment planning for other patients can possibly be developed. For example, a man is not feeling well with his hand. He needs to make an appointment with the clinician who will then arrange a time to take multiple radiographs for assessment. Then he needs to wait for clinical assessment. After some time, he is diagnosed with a hand disease. He now requires a treatment plan for the particular hand disease. During the treatment process, he is required to have his hand radiographs taken for several times until he recovers. The whole process is time-consuming and requires multiple visits to the clinician. What if only a single radiograph of his hand is required

to create the 3D hand for automatic diagnosis and treatment planning. It might only require one extra visit to confirm if he recovers. It would greatly save time and medical resources.

Many research works has been concentrated on registering 3D CT volumes to 2D X-ray images while a few worked on 3D models (see Table 2.1). This thesis concentrates on developing a 3D/2D registration method that aims to register a 3D triangular model of a generic hand to a single radiograph. Our proposed 3D/2D registration method relies on fast X-ray simulation and optimisation algorithms (see Chapter 3). It is not possible to conduct our work in the same setting without the advanced development of computer hardware which enables the real-time X-ray simulation. We start with a proof-of-concept study involves synthetic X-ray images to test feasibility of our method. The performance of our method is assessed using real-world X-ray images from MURA dataset [117] (see Chapter 3). Techniques that aim to improve registration accuracy are investigated (see Chapters 4 and 5). Also, the performance of our method is assessed using another real-world registration problem, which is the registration of a 3D model to a hip radiograph from VHP dataset [1].

1.2 Research hypothesis

How to solve registration problems that involves accurate and fast registration of 3D polygon meshes to X-ray radiographs remains unanswered. We identified there is a lack of research works in this research area. The registration of 3D polygon meshes to X-ray radiographs can be performed accurately and fast using our method which heavily rely on optimisation algorithms and X-ray simulation. With recent advances in graphics processing units (GPUs) and the development of an open-source high-performance X-ray simulation library (gVirtualXRay), 3D/2D registration can be performed fast. Advanced development of optimisation algorithms generally leads to improved accuracy. Synthetic data can be used to test the feasibility of our approach and clinical data can be used to evaluate the performance of our approach in the real world. We summaries questions, and software and research developments with the research hypothesis.

By utilising high-performance X-ray simulation library (gVirtualXRay) and advanced optimisation algorithms, our 3D/2D registration method can be used to register a 3D model to a single 2D image accurately and fast, although care must be given to the choices of metrics and optimisation algorithms which would greatly affect the registration performance.

1.3 Aim and objectives

The aim of this thesis is to develop a 3D/2D registration method that can register a generic 3D hand model to a clinical hand radiograph. The success of this step lays a great foundation for further development of our method in a wider context: any given problems involving the registration of a 3D model to a single 2D image can be solved using our method both accurately and fast.

The objectives are:

- Surveying registration methods (Chapter 2). We review different 3D/2D registration methods and identify what are commonly used input data (3D and 2D) and objective function used in their method.

- Surveying optimisation algorithms and image comparison metrics suitable for registration problems (Chapter 2).

- Proposing a registration method that registers polygon meshes to X-ray radiographs and Examining our registration method in a controlled test case using simulated data (Chapter 3). In this way the feasibility of our method can be assessed.

- Examining our registration method in real-world cases using radiographs with postero-anterior (PA) hand pose (Chapter 3). We assess the performance of our approach and identify area for improvement in our method.

- Addressing deficiencies of the proposed method in real-world cases using different techniques such as image padding and bi-objective optimisation algorithms (Chapter 4).

- Conducting manual registrations and investigating why some images are harder to register (Chapter 5). Further improvement of our method can be achieved by addressing the findings from the manual registration. We assess the performance of our improved method and investigates the relationship between metric pairs in the context of multi-objective optimisation.

- Extending our method to another real-world registration problem which involves a different anatomical region to assess the adaptability of our approach (Chapter 5) to other articulated musculoskeletal anatomy.

1.4 Contributions

- Our proposed registration methods enable the use of fast X-ray simulations in an objective function, i.e. radiographs of 3D models are generated repeatedly to match the real radiograph, which was not possible without advanced development of computer hardware in recent years. Our work is being extended by other researchers (e.g. [159]).

- 3D kinematic configuration of a hand model, i.e. the recovery of the original hand pose in a PA view radiograph. Due to the large variability in clinical practice, no assumption is made in the registration on the source-detector distance (SDD), source-object distance (SOD) and X-ray beam spectrum. The SDD, SOD, rotation angles of all the joints in the hand are automatically estimated by our registration framework. The knowledge of the 3D pose of a hand in a radiograph will allow us to trivially and accurately solve bone segmentation and flag clinically relevant (e.g. rheumatoid) differences from a baseline.

- Registration of a highly articulated and anatomically correct 3D mesh model, hereby referred to as the phantom model, to a real radiograph using a digitally reconstructed

radiograph (DRR) software (https://sourceforge.net/p/gvirtualxray/) [157], [161]. Our proposed optimisation process (inside green box in Figure 3.1) consists of using a single-objective evolutionary algorithm to solve for the articulated 3D pose of the virtual hand that best fits the real radiograph. This experiment was carried out using 15 radiographs from a public dataset, MURA dataset https://stanfordmlgroup.github.io/competitions/mura/ [117].

- Image padding and bi-objective optimisation were used to address deficiencies of our proposed method, i.e. we aim to register images that were not able to registered in previous experiments (Chapter 3). Also, we compared the registration performance between single- and bi-objective optimisation algorithms. This experiment was carried out using 15 radiographs from the MURA dataset.

- We enabled the uses of manual registration to study the causes of problems encountered in earlier experiments (see Chapters 3 and 4) and multi-objective optimisation algorithms to further improve our registration method. The relationship between different metrics in the multi-objective optimisations were investigated. This experiment was carried out using 3 radiographs from the MURA dataset and the hip radiograph from VHP dataset [1].

1.5 Outline of the thesis

This thesis is divided into seven chapters. The first chapter discusses the motivation of our researches, presents aims and objectives, and lists main contributions to science as well as our published papers. The second chapter introduces the general concept of 3D/2D registration, discusses traditional 3D/2D registration methods, and lists common metrics and optimisation algorithms. The third chapter proposes a 3D/2D registration method, tests the feasibility of our approach and compares the performance between different optimisation algorithms. Also, the performance of our registration method is assessed in real-world registration problems. The fourth chapter investigates the effect of image padding in 3D/2D registration and addresses deficiencies of the proposed method using a bi-objective optimisation algorithm. The fifth

chapter investigates deficiencies of the proposed method using manual registrations, improves registration accuracy using multi-objective optimisation algorithms and analyses the relationship between different metrics in multi-objective optimisation. Our approach is then adapted to another anatomical region, the hip, which is to test the adaptability of our method. The last chapter concludes our research work and discusses the limitation of our approaches, and introduces further research directions.

Chapter 2

Related work

2.1 Introduction

Registration is the process of transforming one coordinate system to match with another coordinate system. Solving an image registration problem typically involves the transformation of a moving dataset (called *source*) to match a fixed dataset (called *target*). The transform could be rigid or non-rigid. While the rigid transform only allows rotation and translation, non-rigid transform allows free-form mappings but typically imposes some restrictions [177].

3D/2D registration methods concentrate on the registration of the *source* (e.g. a CT volume) to the *target* (e.g. a 2D image). Figure 2.1 illustrates the general concept of establishing the dimensional correspondence in traditional 3D/2D registration methods. There are four main components in traditional 3D/2D registration method: i) inputs (the *source* and the *target*), ii) the transform, iii) the objective function and iv) the optimisation algorithm. Changes to each individual component would result in different performance of the registration methods. Depending on the registration process, 3D/2D registration methods could be classified as *extrinsic*, *intrinsic* and *calibration-based* [92]. More details will be discussed in Section 2.2.

Machine learning (ML) is a very popular research topic where deep learning (DL) is the leader in this field. DL has been widely used in solving complex problems such as image classification [79] and speech recognition [53]. In recent years, deep learning has been applied

in solving 3D/2D medical registration problems where the registration of the CT volumes (or a 3D model) to 2D X-ray images is performed [24], [39], [57], [172], [179].

3D/2D registration has many clinical applications such as image-guided radiation therapy (IGRT) [17], [31], [36], [61], [116], image-guided radiosurgery (IGRS) [22], [41], [112], [113], [122], [123], image-guided endoscopy (IGE) [93], [134] and motion correction [21], [35], [46], [72], [88], [90].

The rest of the chapter is organised as follow: Section 2.2 summaries different 3D/2D registration methods and shows what type of inputs is used in traditional 3D/2D registration methods, Section 2.3 discusses some image registration methods that utilise deep learning techniques, Section 2.4 presents similarity and dissimilarity measurements that are used in the objective function in 3D/2D registration methods, Section 2.5 lists different types of optimisation algorithms in 3D/2D registration methods and Section 2.6 presents basic information about X-ray simulation and a library that is used to generate X-ray images.

2.2 Traditional 3D/2D registration methods

In all registration methods, the dimensional correspondence between the *source* and the *target* must be established. There are three strategies that can be used to achieve this goal in traditional 3D/2D registration methods: i) projection, ii) back-projection and iii) reconstruction (Figure 2.1).

Figure 2.2 shows registration methods based on the projection strategy. *Source* is transformed and then projected onto a 2D plane. In this way 3D/2D registration becomes 2D/2D. The distance between the 2D projections of the *source* and the *target* is then minimised with an optimisation algorithm to obtain optimal solutions. Figure 2.3 shows registration methods based on back-projection strategy. The *target* is back-projected into the 3D space. The best solution is found by optimising the distance between the transform of the *source* and the back-projected 3D model (or 3D gradient vectors). The reconstruction strategy is similar to the back-projection

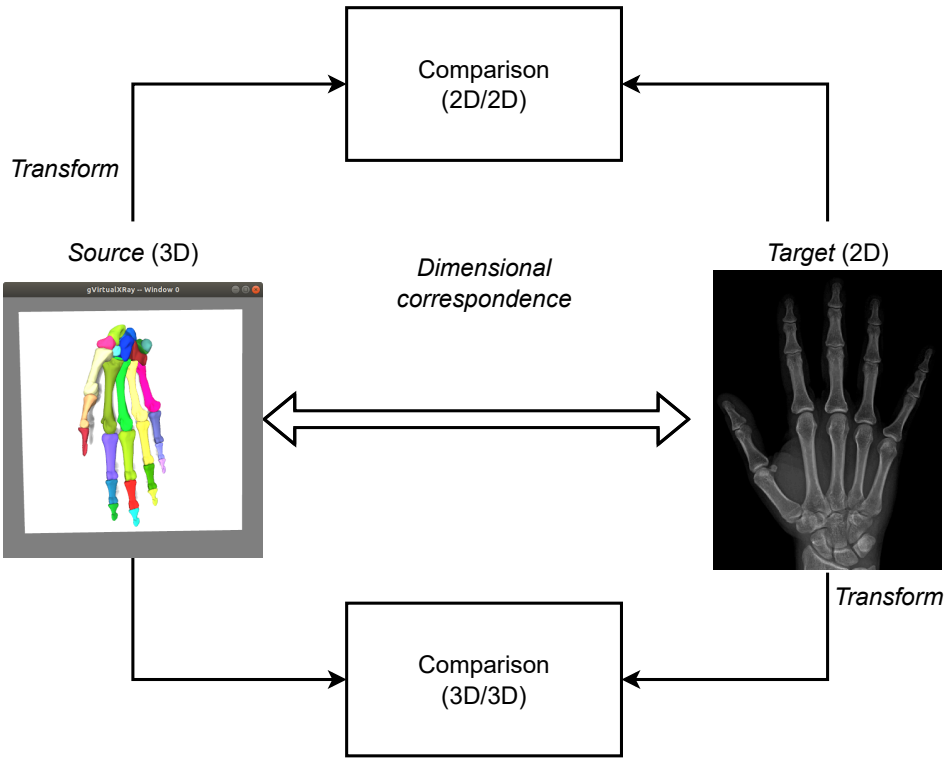


Figure 2.1: *General strategies of establishing the dimensional correspondence in traditional 3D/2D registration methods. Examples of 3D and 2D data are shown including a 3D hand model and an X-ray images. The 3D Source can be projected onto a 2D plane where comparison is performed in 2D space. The 2D Target can be back-projected or reconstructed into 3D space where comparison is performed in 3D space.*

strategy where 3D/2D registration becomes 3D/3D. Usually, the *target* consists of multiple 2D images (or 2D gradient images) which are then reconstructed to become a 3D model (or 3D images, or 3D gradient vectors).

2.2.1 *Extrinsic* methods

Extrinsic methods often require the assist of artificial objects such as stereotactic frames [65] and fiducial markers [6], [69], [85], [136], [138], [145]. Markers can be implanted into the human body (e.g. as shown in Figure 2.5). All three strategies can be used to establish the dimensional correspondence in *extrinsic* registration methods (Figures 2.2, 2.3 and 2.4). For example, the projection strategy was used in [69], [136], [138]. Firstly, a fiducial marker that is made of gold is inserted into the tumour or placed near the tumour. The position of the marker in the CT image is projected on onto the 2D plane. Registration is performed by minimising

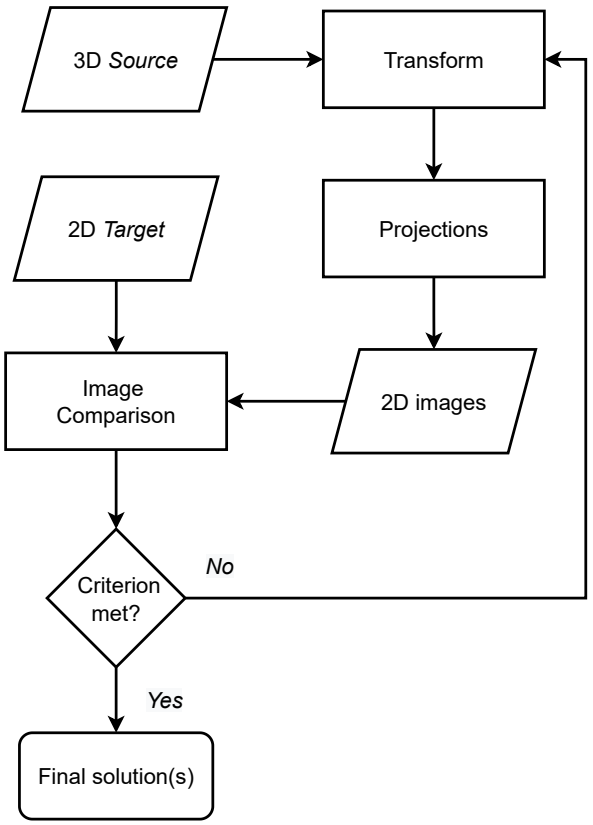


Figure 2.2: *A diagram of 3D/2D registration methods that utilise the projection strategy. The distance is minimised between 2D projections of the source and the 2D target.*

the distance between projected marker from the CT image and the X-ray image. Tang, Ellis and Fichtinger [145] proposed a 3D/2D registration method that based on the back-projection strategy. Virtual X-rays were firstly created to connect markers on 2D images with the X-ray source. Then the registration was performed by minimising distance between back-projected markers and virtual X-rays. The reconstruction strategy was applied in [6], [85]. The 3D markers were reconstructed from multiple radiographs where the position of 3D markers were determined. The registration was performed by minimising the distance between reconstructed 3D markers and reference 3D markers that are obtained from the CT dataset.

2.2.2 *Intrinsic* methods

Intrinsic methods concentrate on anatomical structures in the data and unlike *extrinsic* methods, in the sense that they solely rely on information contained in the data. Depending on what

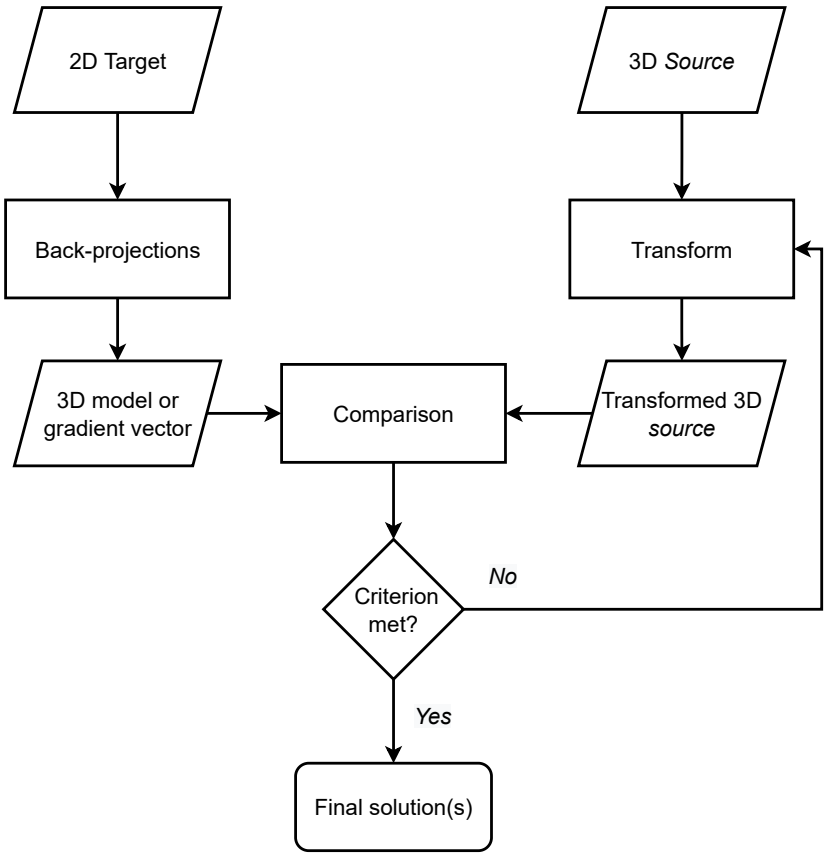


Figure 2.3: *A diagram of 3D/2D registration methods that utilise the back-projection strategy. The distance is minimised between the 3D source and the back-projected 3D model (or 3D gradient vectors) from the target.*

information is determined to be important, *Intrinsic* methods can be classified as *feature-*, *intensity-* and *gradient-based* [92]. *Feature-based* methods concentrate on features such as points, a set of points that are forming a line, curve, contour or surface. While *intensity-based* methods rely on information that is contained in 3D voxels and 2D pixels, *gradient-based* methods rely on information that is contained in 3D or 2D gradient vectors.

2.2.2.1 *Feature-based* methods

Feature-based 3D/2D registration methods are looking to find correspondences between 3D features and 2D features that are extracted from 3D and 2D dataset by the use of image segmentation [52]. Image segmentation is a technique that separates the data into different regions where complexity of data is greatly reduced. A commonly used image segmentation

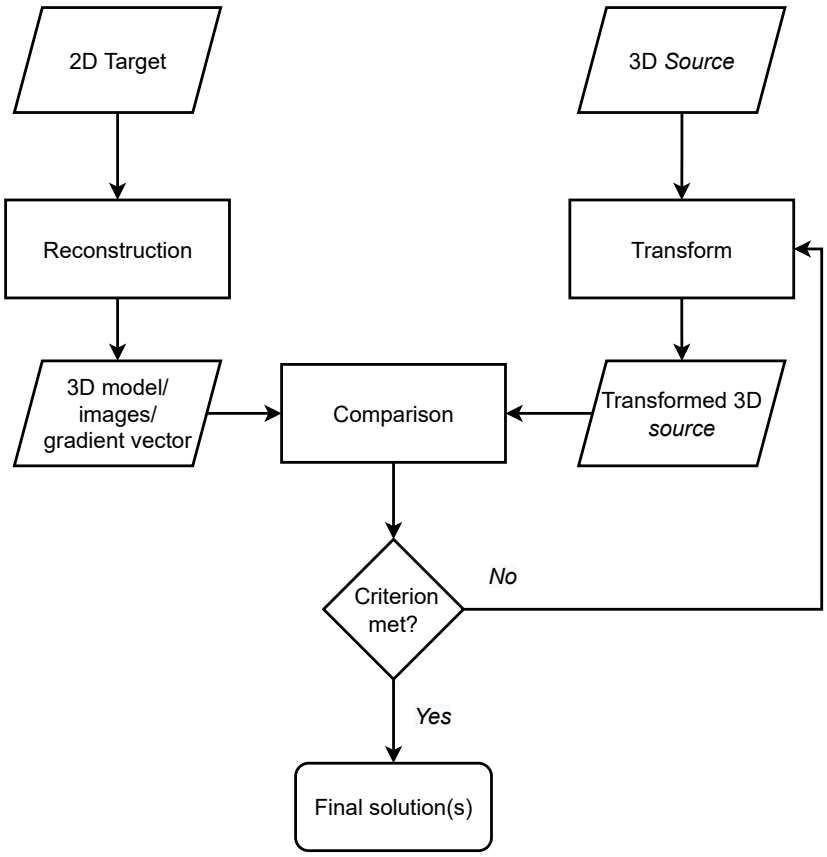


Figure 2.4: *A diagram of 3D/2D registration methods that utilise the reconstruction strategy. The distance is minimised between 3D source and the reconstructed 3D model (or 3D image, 3D gradient vectors) from the target.*

technique is *thresholding*. The simplest form of *thresholding* approach is binary where pixels are separated into two classes based on a threshold value. A more complex *thresholding* method which partitions the data into many different classes was developed in [5]. Other segmentation techniques are also developed such as *clustering* [26] and *neural networks* [45]. However, segmentation consumes a lot of computational resources (or human resources if manual segmentation is performed) and inaccurate segmentation would result in reduced registration accuracy. Depending on the segmentation methods that are used, reduced data might also lead to loss of some important features. Some methods are developed to reduce the effect of inaccurate segmentation [76], [173].

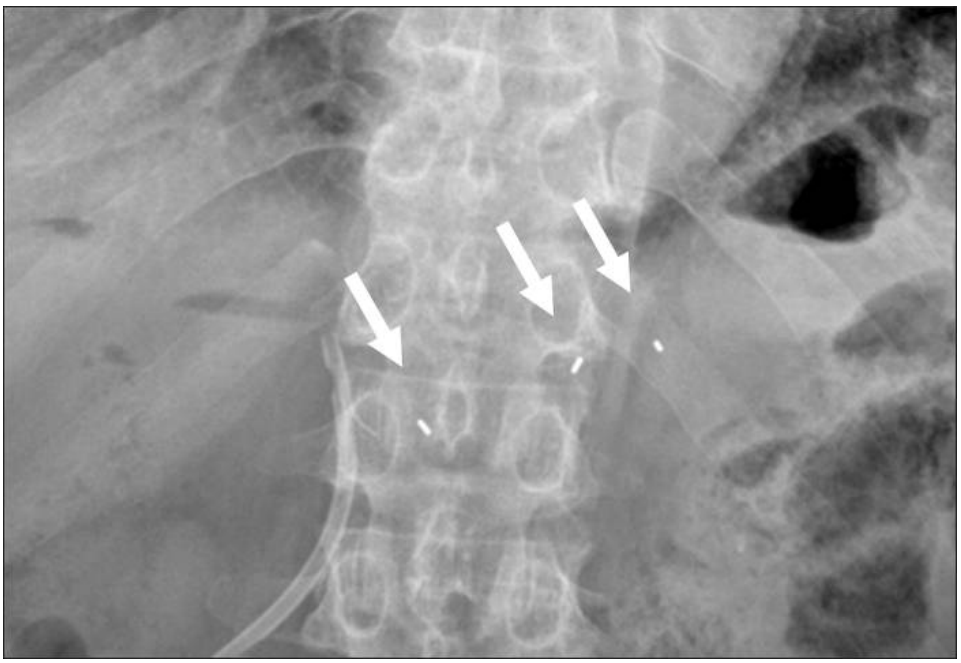


Figure 2.5: *An example radiograph of abdomen that contains three cylindrical fiducial markers. Source: Figure 2 in [23].*

Depending on different targeting correspondence pairs between 3D and 2D features, feature-based 3D/2D registration methods can be classified as *point-to-point (PTP)*, *curve-to-curve (CTC)* and *surface-to-curve (STC)*.

PTP methods aim to locate the position of 2D points in a 2D image and corresponding 3D points in 3D image. The choice of points' location is usually determined manually by an operator. For example, corresponding point pairs were marked manually and registration was performed using an *extrinsic* registration method as presented in [14]. PTP method is time-consuming and some important features may lost due to human intervention. However, it is fast and easy to implement where in some case, it could be used for an initial approximation in a more complex registration method [10].

CTC methods aim to locate curves in 3D and 2D dataset. A *CTC* registration method based on the projection strategy was developed in [50]. The registration was performed by minimising the distance between the projected 3D curve and the 2D curve. The back-projection strategy was applied in [37]. 2D curve are back-projected into 3D space and the registration is performed by

minimising the distance between 3D curve and back-projected 2D curve. The 3D/2D registration becomes 3D/3D where curve correspondences are established.

STC methods aim to determine correspondences between surface and curve, from the 3D and the 2D dataset, respectively. A *STC* registration method was presented in [181] where both projection and reconstruction strategy were applied. A 3D point distribution model (PDM) was created and was projected onto the 2D plane. 2D point pairs were obtained by minimising the distance between projected point sets and 2D points sets in X-ray images. Then 2D point pairs were then reconstructed into 3D space.

2.2.2.2 *Gradient-based* methods

Gradient-based methods aim to find the correspondence between 3D and 2D gradient vectors that are extracted from 3D and 2D dataset, which were firstly introduced in [86], [148]. All three strategies can be used to establish dimensional correspondence (Figures 2.2, 2.3 and 2.4). Tomazevic, Likar, Slivnik *et al.* [148] proposed a *gradient-based* 3D/2D registration method that registers a 3D CT or Magnetic Resonance (MR) image to 2D X-ray images. Dimensional correspondence is established by applying the back-projection strategy. The 3D gradient vector field was extracted from a 3D CT or MR image and 2D gradient vectors were extracted from 2D X-ray images. Registration is performed by back-projecting 2D gradient vectors in 3D where the distance (both magnitude and direction) between the 3D gradient vector field and 3D back-projecting gradient vectors is minimised. A *gradient-based* method based on the projection strategy was developed in [167] where registration is performed by maximising the similarity between the projection of the 3D gradient image and the 2D gradient image. Another *gradient-based* method that relys on the reconstruction strategy was presented in [91]. The 3D gradient vector field and 2D gradient vectors were extracted from a 3D CT (or MR image) and 2D X-ray images, respectively. Multiple 2D gradient vectors were used to reconstruct a 3D gradient vector field. Registration is performed by maximising the similarity between two 3D gradient vector fields.

2.2.2.3 *Intensity-based* methods

Most *intensity-based* methods are *digitally reconstructed radiograph (DRR)-based* where registration is performed using DRRs that are generated from the 3D dataset and 2D X-ray images. A DRR is a radiographic image that is generated from a CT volume. *DRR-based* registration methods rely on similarity or dissimilarity measurements between the generated DRR and the 2D X-ray image. More details about similarity or dissimilarity metrics will be discussed in Section 2.4 (also in Table 2.2).

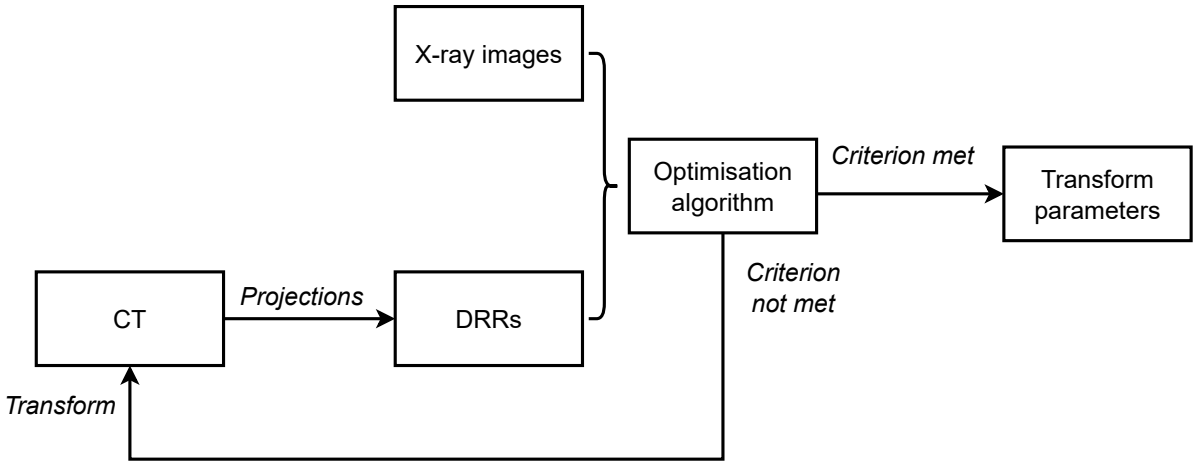


Figure 2.6: *A DRR-based 3D/2D registration method that register a CT volume to X-ray images where the projection strategy is applied by Gong, Stewart and Abolmaesumi [47].*

Table 2.1 gives an overview of inputs that are required in different *intensity-based* 3D/2D registration methods. Most methods are concentrating on registering a 3D CT volume to X-ray images so that the projection strategy is used to achieve the dimensional correspondence. For example, Gong, Stewart and Abolmaesumi [47] proposed an approach to register a CT volume to multiple X-ray images with different views (Figure 2.6). Multiple DRRs are generated from CT data by rigid transformation (rotation and translation). The similarity measurements between DRR and X-ray image pairs are then optimised which gives a set of metric values. They argued that information is lost if all metric values are combined into a single metric value. This is due to the fact that multiple X-ray images with different views are used. Each metric value in the metric value sets is different from each other. In other words, some metric values

would impose more weights on the determination of the final metric value than if a single metric value was used.

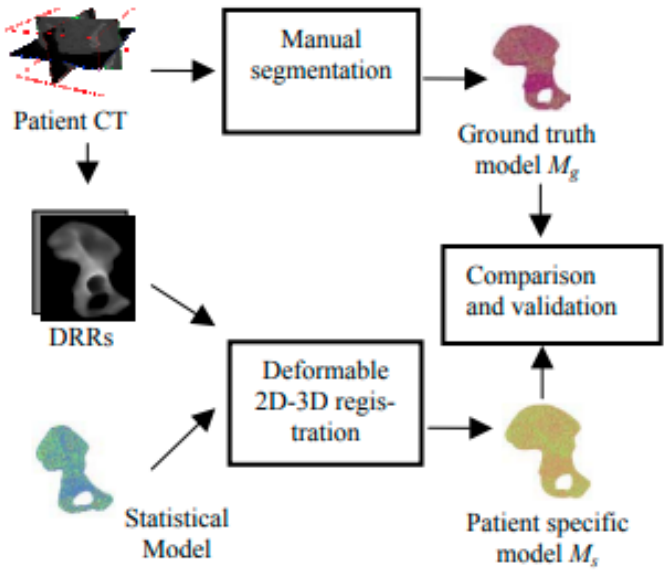


Figure 2.7: *A DRR-based 3D/2D registration method that registers a 3D model to X-ray images where the projection strategy is applied by Yao and Taylor (Source: Figure 1. Experiment flow chart [175]). The 3D model is reconstructed using DRRs from a CT image.*

A similar registration approach that utilises the projection strategy is proposed by Yao and Taylor [175]. Their approach aims to register a 3D statistical model to multiple X-ray images (Figure 2.7). Firstly, a 3D statistical density model is reconstructed based on contours of the object from CT images. DRRs are obtained by non-rigid transformation of the 3D statistical density model. Registration is performed by maximising the similarity measurements between the DRRs that are generated from the 3D statistical density model and the DRRs that are generated from a CT dataset. They have identified some factors that might affect the accuracy of registration:

- **Number of 2D projection views**. Generally speaking, when number of views increases, the registration accuracy improves although improvements are not significant when number of views is greater than three.

- **The angle between two projection views**. The accuracy is the highest when two projections are orthogonal to each other and the accuracy is the lowest when the angle between two projections is 0 degree.

- **Image noise**. When the noise level (magnitude) is high, the registration accuracy is reduced.

- **Image distortion**. Spatial distortion on X-ray images reduces the accuracy of registration.

- **Co-registration error**. Registration accuracy decreases when co-registration of DRRs produces error. Co-registration is the process of registering multiple 2D images with different views (i.e. radiographs taken at different angles) to determine their relative pose in relation to a fixed coordinate system.

Other *DRR-based* registration methods where the 3D model that is represented by tetrahedral mesh were also developed in [58], [127], [128] and different mesh representation (triangular mesh) also explored [4], [154]. The related literature can also be found in Table 2.1.

We discussed some *DRR-based* registration methods that utilise the projection strategy where 3D/2D registration becomes multiple 2D/2D registrations. The reconstruction strategy can also be applied. For example, Tomazevic, Likar and Pernus [147] proposed a registration method that is based on the reconstruction strategy. Their approach aims to register 3D images to 2D X-ray images. The 3D image is reconstructed from multiple 2D X-ray images. Registration is performed by optimising the similarity measurements between reconstructed 3D image and the reference 3D image. They argued that valuable information is lost and registration accuracy might be reduced in *DRR-based* registration methods where the projection strategy is applied.

2.2.3 *Calibration-based* methods

Calibration-based methods aim to establish dimensional correspondence by incorporating an imaging system which combines both 2D and 3D data. For example, Rhode, Hill, Edwards

Table 2.1: *Different 3D and 2D data that are used in intensity-based 3D/2D registration methods*

Related literature	2D data	3D data
[15], [16], [18], [25], [30], [33], [42], [47], [54], [59], [68], [73]–[75], [77], [80], [84], [100], [101], [104], [109]–[111], [120], [121], [124]–[126], [154], [164]–[166], [178], [180], [182], [183]	X-ray	CT
[70]	DSA	Angiography
[4], [81]	X-ray	Model (triangular mesh)
[58], [127], [128], [175]	X-ray	Model (tetrahedral mesh)

et al. [118] proposed a registration method that registers MR images to X-ray images. Their approach relies on X-ray and MR interventional suite (XMR) systems where patients can be transferred between MR and X-ray systems. Imaging devices are firstly calibrated and a tracking system is used to determine the relative position of X-ray set, patient and the MR scanner. The tracking information enables the registration of projections that are generated from the MR image to X-ray image. Also, registration of CT to digital subtracted angiograph (DSA) can also be performed by using the *calibration-based* registration method [60].

2.3 *Deep learning-based* registration methods

Early development of *DL-based* registration methods concentrates on incorporating deep learning techniques into existing registration methods such as *intensity-based* registration methods. For example, Wu *et al.* [171] proposed a registration framework where feature extraction of 3D MR brain images is performed using a stacked two-layer convolutional network. Then registration is performed by optimising the similarity metric between two sets of features. Simonovsky *et al.* [139] developed a similarity metric based on a convolutional neural network (CNN) which is then applied into the optimisation process.

In recent years, registration methods that aim to establish correspondences between input data (the *target* and the *source*) are also developed. Miao *et al.* [96] proposed a rigid registration

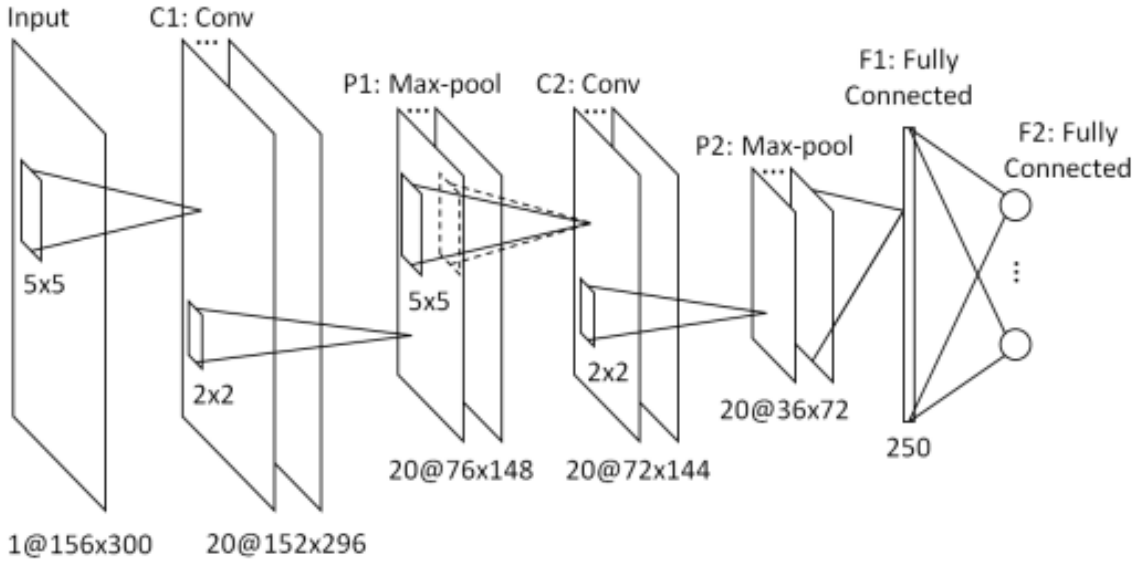


Figure 2.8: *The CNN architecture used by Miao et al. Source: Figure 4 in [96].*

method which aims to match DRRs of 3D model with X-ray images using CNN. Feature extraction is performed on original images and residual differences between extracted images are computed. The difference image is then fed to a six-layer CNN (as shown in Figure 2.8) where the output layer (F2) corresponds to transformation parameters. Their model is trained using synthetic images and validation is carried out using a dataset for Virtual Implant Planning System (VIPS) [155]. Results show that their method outperforms some intensity-based method such as mutual information (MI) and gradient correlation (GC) both in terms of accuracy and speed. Similar approaches that concentrate on rigid registration using CNN can be found in [129], [141].

Cao *et al.* [20] proposed a deformable registration framework based on CNN which the complex mappings between input pairs and the deformation field can be found. Their approach is shown in Figure 2.9. They firstly extract image patches from inputs at same location. Those image patches are then fed into a CNN to compute displacement vectors. A series of displacement vectors constitute a deformation field which can be used to perform registration. Also, similarity measurements between those image patches are used to guide the learning process. They used the active-points guided sampling strategy which prefers images patches with higher gradient

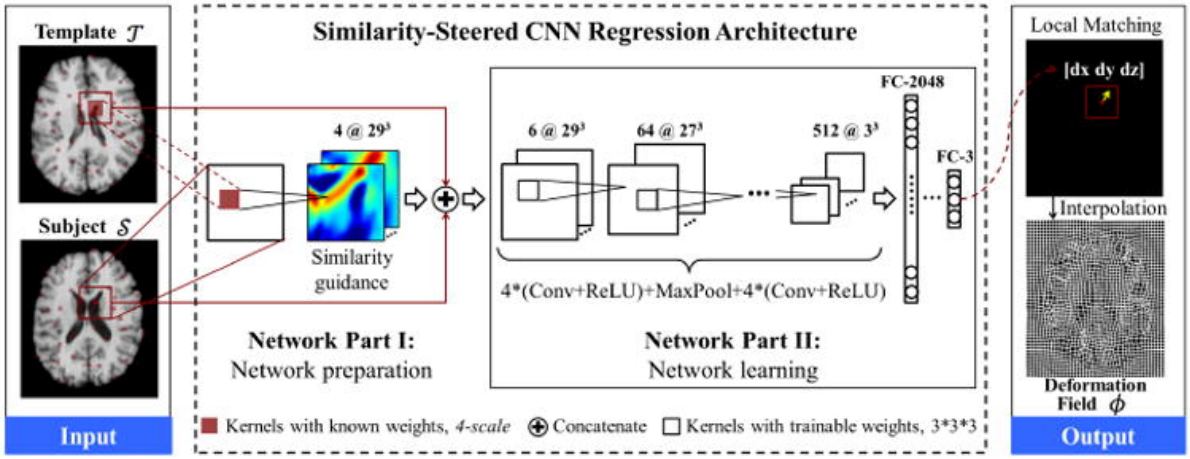


Figure 2.9: *An input-to-deformation mapping method proposed by Cao et al. Source: Figure 1 in [20].*

magnitude and displacement values, i.e. voxels with rich anatomical information would be more likely to be sampled. Two datasets contains MR brain volumes are used to evaluate their method including the LONI Probabilistic Brain Atlas (LPBA40) [137] and Alzheimer's Disease Neuroimaging Initiative (ADNI) [98]. Although 2D images are used for illustration (as shown in Figure 2.9), their evaluation is conducted using MR brain volumes. The registration results show that their approach outperforms some state-of-the-art methods such as SyN [7] and Demons [153]. The area with better alignment is highlighted by yellow arrows and lines in Figure 2.10.

Lv *et al.* [88] also developed a deformable registration method based on CNN where the registration of abdominal MR images is performed to correct respiratory motion. The input image pairs are fed into CNN which outputs the displacement vector field where registration is performed. They used self-acquired dataset for evaluating their method. Their approach performs well in terms of evaluation metrics and visual assessment by two radiologists. Other research works that are dedicated to deformable registration methods can be found in [119], [142], [150].

The first main challenge that *DL-based* 3D/2D registration methods faces is the requirement of a large number of annotated training data. However, those dataset might not be easily accessible.

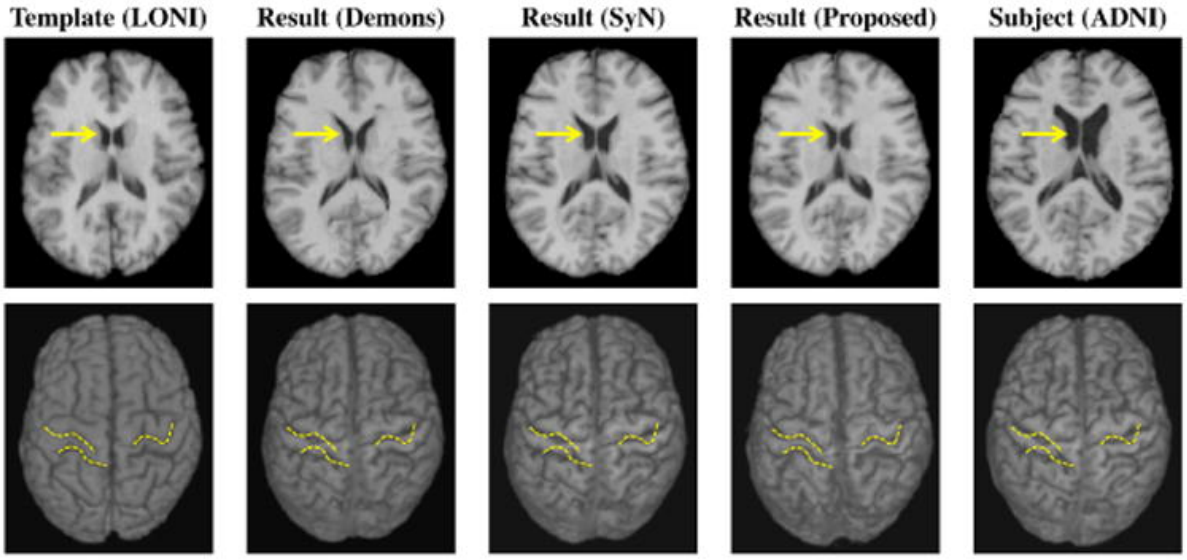


Figure 2.10: *The registration result of their proposed method by Cao et al. Source: Figure 5 in [20].*

Also, the training process of the learning model is often time-consuming where advanced hardware such as high-performance GPU is required, which might not be feasible in some cases.

2.4 Survey of similarity and dissimilarity metrics

Some commonly used similarity and dissimilarity metrics in 2D/3D registration are discussed in this section. Table 2.2 displays those metrics along with their equation and related literature. We define two images, the target image $\mathbf{Y}$ and the predicted image $\hat{\mathbf{Y}}$, with same number of pixels along the horizontal (w) and vertical axis (h). The standard deviations of the pixel values in target and predicted images are represented by $\sigma_{\mathbf{Y}}$ and $\sigma_{\hat{\mathbf{Y}}}$. $\sigma^2_{\hat{\mathbf{Y}}}$ are variance of $\mathbf{Y}$ and $\hat{\mathbf{Y}}$.

2.4.1 Normalised cross correlation

Normalised cross correlation (NCC) [15] is a similarity metric. The NCC between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\mathrm{NCC}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \frac{\hat{\mathbf{Y}}(i,j)\mathbf{Y}(i,j)}{\sigma_{\hat{\mathbf{Y}}}\sigma_{\mathbf{Y}}} \tag{2.1}$$

Table 2.2: *Summary of metrics that are used in 3D/2D intensity-based registration methods*

Metrics	Equation	Related literature	Number count
NCC	Eq. 2.1	[15], [25], [30], [47], [54], [63], [70], [71], [73], [74], [77], [80], [83], [99]–[101], [111], [125], [130], [178]	20
MI	Eq. 2.4	[25], [33], [54], [71], [73]–[75], [84], [99], [111], [121], [124]–[126], [130], [180], [182], [183]	18
PI	Eq. 2.7	[15], [16], [33], [54], [55], [71], [74], [84], [109], [111], [120], [125], [140], [164]–[166]	16
GD	Eq. 2.8	[19], [54], [59], [71], [74], [94], [105], [109]–[111], [125]	11
GC	Eq. 2.9	[18], [30], [54], [71], [83], [104], [146]	7
SSD	Eq. 2.10	[42], [63], [68], [71], [180], [182]	6
EOD	Eq. 2.12	[54], [74], [111], [125]	4
SNLC	Eq. 2.14	[71], [77], [81], [82]	4
CS	Eq. 2.15	[25], [102], [103], [130]	4
CR	Eq. 2.16	[25]	1
SRC	Eq. 2.17	[15]	1

Zero-mean normalised cross correlation (ZNCC) is a similarity metric and ZNCC between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defines as:

$$\text{ZNCC}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \frac{\left(\hat{\mathbf{Y}}(i,j) - \overline{\hat{\mathbf{Y}}}\right)\left(\mathbf{Y}(i,j) - \overline{\mathbf{Y}}\right)}{\sigma_{\hat{\mathbf{Y}}}\sigma_{\mathbf{Y}}} \tag{2.2}$$

$\overline{\mathbf{Y}}$ and $\overline{\hat{\mathbf{Y}}}$ are mean pixel value of target and predicted images. ZNCC is easy to interpret. It ranges from -1 to 1, where i) the value is close to 1, the two images are highly similar which implies high level of correlation, ii) the value is 0, two images are extremely different which implies there is no correlation, iii) the value is -1, one image is the negative of the other image which implies they are anti-correlated or inversely correlated. ZNCC is widely used in image comparison due to its robustness in template matching [34], [78]: it is robust to linear transformations, i.e. to gain and bias differences. It means that if $\hat{\mathbf{Y}} = a\mathbf{Y} + b$, ZNCC is 1 when $a > 0$ and -1 if $a < 0$.

2.4.2 Mutual Information

MI compares the statistical information between two images [143]. MI measures of the mutual dependence between the two random variables and can be used as a similarity metric between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$. It is defined as:

$$\mathrm{MI}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \mathrm{H}(\hat{\mathbf{Y}}) - \mathrm{H}(\hat{\mathbf{Y}}|\mathbf{Y}) \tag{2.3}$$

Where $\mathrm{H}(\hat{\mathbf{Y}})$ is entropy for $\hat{\mathbf{Y}}$ and $\mathrm{H}(\hat{\mathbf{Y}}|\mathbf{Y})$ is the conditional entropy for $\hat{\mathbf{Y}}$ given $\mathbf{Y}$. MI is used to detect non-linear relationship between images. The value of MI is non-negative and is often hard to interpret. It can be written in the normalisation form:

$$\mathrm{NMI}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{\mathrm{H}(\hat{\mathbf{Y}}) + \mathrm{H}(\mathbf{Y})}{\mathrm{H}(\hat{\mathbf{Y}}, \mathbf{Y})} \tag{2.4}$$

Where $\mathrm{H}(\mathbf{Y})$ and $\mathrm{H}(\hat{\mathbf{Y}})$ are the entropy for $\mathbf{Y}$ and $\hat{\mathbf{Y}}$, respectively.

$$\mathrm{H}(\mathbf{Y}) = -\frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \mathbf{Y}(i, j) \log \frac{1}{\mathbf{Y}(i, j)} \tag{2.5}$$

And $\mathrm{H}(\hat{\mathbf{Y}}, \mathbf{Y})$ is the joint entropy (Equation 2.6).

$$\mathrm{H}(\hat{\mathbf{Y}}, \mathbf{Y}) = -\frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \mathbf{Y}(i, j)\hat{\mathbf{Y}}(i, j) \log \frac{1}{\mathbf{Y}(i, j)\hat{\mathbf{Y}}(i, j)} \tag{2.6}$$

The value of NMI is also easy to interpret. It ranges from 1 (uncorrelated) to 2 (perfectly correlated).

2.4.3 Pattern intensity

Pattern intensity (PI) is a similarity metric which measures the changes of structural information on the difference image [165]. A small value is assigned to area within the neighbourhood of structure such as edges and lines while a large value is assigned to area with little gray-value variation. PI is defined as:

$$\mathrm{PI}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \sum_{i,j} \sum_{(i-v)^2+(j-w)^2 \leq r^2} \frac{\sigma^2}{\sigma^2 + (\Delta\mathbf{Y}(i,j) - \Delta\mathbf{I}(v,w))^2} \tag{2.7}$$

r and σ are two parameters set by the user. r is used to define the size of area that will be considered and σ is the sensitivity that defines the area to be structure or not. $\Delta\mathbf{Y}$ is the difference image as defined in Equation 2.13. $\Delta\mathbf{I}$ is the difference image of selected regions defined by r.

2.4.4 Gradient difference

Gradient difference (GD) [111] is a dissimilarity (or distance) metric and GD between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\mathrm{GD}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \left\{ \frac{\sigma_{\mathbf{Y}}^2}{\sigma_{\mathbf{Y}}^2 + \Delta \frac{\partial \mathbf{Y}}{\partial i}^2} + \frac{\sigma_{\hat{\mathbf{Y}}}^2}{\sigma_{\hat{\mathbf{Y}}}^2 + \Delta \frac{\partial \hat{\mathbf{Y}}}{\partial j}^2} \right\} \tag{2.8}$$

The value of GD is non-negative and the best value can be achieved is 1 where two gradient images are same.

2.4.5 Gradient correlation

GC [111] is a similarity metric and it is the average value of Normalised cross correlation (NCC) between gradient images, where GC between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as follow:

$$\mathrm{GC}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{1}{2}\left\{NCC\left(\frac{\partial \mathbf{Y}}{\partial i}, \frac{\partial \hat{\mathbf{Y}}}{\partial i}\right) + NCC\left(\frac{\partial \mathbf{Y}}{\partial j}, \frac{\partial \hat{\mathbf{Y}}}{\partial j}\right)\right\} \tag{2.9}$$

Where $\frac{\partial}{\partial I}$ is calculated using Sobel filter. The Sobel filter, or Sobel operator is a technique to find approximations of the derivatives which is achieve by convolving the a matrix (kernel with size of n by n) with the original image.

2.4.6 Sum or Mean squared error

Sum of squared error (SSE), also known as sum of squared distance (SSD) is a dissimilarity (or distance) metric and SSE between [42] between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\mathrm{SSE}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \sum_{j}^{h}\sum_{i}^{w}\left(\mathbf{Y}(i,j) - \hat{\mathbf{Y}}(i,j)\right)^2 \tag{2.10}$$

Mean square error (MSE) is a dissimilarity (or distance) metric and it is calculated by dividing SSE with the total number of pixels, denoted by $w \times h$.

$$\mathrm{MSE}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{1}{w \times h} \times SSE \tag{2.11}$$

where w and h are the number of pixels in $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ along the horizontal and vertical axis respectively. The best values of SSE and MSE are 0 where two images are same. MSE tends to penalise outliers, i.e. a larger value is assigned to the outlier point. On the other hand, MSE tends to be unreliable if there is a small number of data points.

2.4.7 Entropy of difference

Entropy of difference (EOD) [74] is a dissimilarity (or distance) metric, which measures entropy on the difference image $\Delta\mathbf{Y}$, is defined as:

$$\mathrm{EOD}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = -\frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \Delta\mathbf{Y}(i,j) \log \Delta\mathbf{Y}(i,j) \tag{2.12}$$

where w and h are the number of pixels in the difference image $\Delta\mathbf{Y}$.

$$\Delta\mathbf{Y} = \mathbf{Y} - s \times \hat{\mathbf{Y}} \tag{2.13}$$

Where s is a scaling factor. Normally, $s = 1$ when images have same size and intensity range. The best value of EOD is 0 where two images are same.

2.4.8 Sum of local normalised correlation

Sum of local normalised correlation (SLNC) [82] is a similarity metric and SLNC between $\hat{\mathbf{Y}}$ and $\mathbf{Y}$ is defined as:

$$\mathrm{SLNC}\left(\hat{\mathbf{Y}}, \mathbf{Y}\right) = \frac{1}{|N|} \sum_{n \in N} NCC(\hat{\mathbf{Y}}, \mathbf{Y}, \mathbf{P}_n) \tag{2.14}$$

Where $NCC(\hat{\mathbf{Y}}, \mathbf{Y}, \mathbf{P}_n)$ is the normalised correlation coefficient between two images at area $\mathbf{P}_n$. N is the total number of data points.

2.4.9 Chi-square

Chi-square (CS) [25] is a dissimilarity metric which measures the distance between two images $\hat{\mathbf{Y}}$, $\mathbf{Y}$ and it is defined as:

$$\mathrm{CS}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \frac{(\mathbf{Y}(i,j) - \hat{\mathbf{Y}}(i,j))^2}{\mathbf{Y}(i,j) + \hat{\mathbf{Y}}(i,j)} \tag{2.15}$$

2.4.10 Correlation ratio

Correlation ratio (CR) [25] is a similarity metric and CR between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\mathrm{CS}(\mathbf{Y}, \hat{\mathbf{Y}}) = 1 - \frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \frac{(\mathbf{Y}(i,j) - \hat{\mathbf{Y}}(i,j))^2}{\sigma_{\mathbf{Y}}^2} \tag{2.16}$$

2.4.11 Stochastic rank correlation

Stochastic rank correlation (SRC) [15] is a similarity metric and it is based on Spearman's rank correlation coefficient which assesses the rank correlation between two variables. Standard RC is computed using total number of pixels (n) but SRC concentrates on total number of pixels in the selected region (R). SRC between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\mathrm{SRC}(\mathbf{Y}, \hat{\mathbf{Y}}) = 6 \times \sum_{r}^{R} \frac{(\mathbf{Y}_{\rho_r} - \hat{\mathbf{Y}}_{\rho_r})^2}{r(r^2 - 1)} \tag{2.17}$$

Where $r \ll N$ and ρ_r is the set of ranks of pixels in the selected region R.

The rank correlation (RC) between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\mathrm{RC}(\mathbf{Y}, \hat{\mathbf{Y}}) = 6 \times \sum_{n}^{N} \frac{(\mathbf{Y}_{\rho_n} - \hat{\mathbf{Y}}_{\rho_n})^2}{n(n^2 - 1)} \tag{2.18}$$

2.4.12 Mean absolute error

Mean absolute error (MAE) is a dissimilarity (or distance) metric. The MAE between two images $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ is defined as:

$$\text{MAE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{w \times h} \sum_{j}^{h} \sum_{i}^{w} \left| \hat{\mathbf{Y}}(i, j) - \mathbf{Y}(i, j) \right| \tag{2.19}$$

where w and h are the number of pixels in $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ along the horizontal and vertical axis respectively. MAE is always non-negative and the best value can be achieved is 0. MAE is often easy to interpret since it has same units of measurement as the data. Also, MAE is less sensitive to outliers which makes it more suitable for assessing the average model performance [170] than MSE.

2.4.13 Mean absolute percentage error

Mean absolute percentage error (MAPE) is also known as relative error is also popular to quantify dissimilarities as it can be expressed as a percentage.

$$\text{MAPE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{100\%}{w \times h} \sum_{y}^{h} \sum_{x}^{w} \left| \frac{\mathbf{Y}(i, j) - \hat{\mathbf{Y}}(i, j)}{\mathbf{Y}(i, j)} \right| \tag{2.20}$$

However, MAPE will be excessively high and unreliable when $\mathbf{Y}(i, j)$ is close to 0.

2.4.14 Practical examples

The similarity or dissimilarity metrics are widely used for comparing two images. However, the metric value does not always tell the whole story, i.e. sometimes the metric value can be misleading. For example, Figure 2.11 shows three different images with same size defined by width and height. The question is which image (2 or 3) is more similar to Image 1. At first glance Image 2 appears to be more similar to Image 1 visually. Let's take two metrics (MAE and RMSE) for statistical comparisons. The MAE between Image 1 and 2 is 0.2 and MAE

is 0.22 between Image 1 and 3, where it implies that Image 1 and 2 are more similar to each other. However, RMSE between Image 1 and 2 is 0.63 while RMSE between Image 1 and 3 is 0.6. Image 3 is more similar to Image 1 when RMSE is used for comparison. For both image, ZNCC equals to zero which indicates that neither Image 2 nor 3 is similar to Image 1.

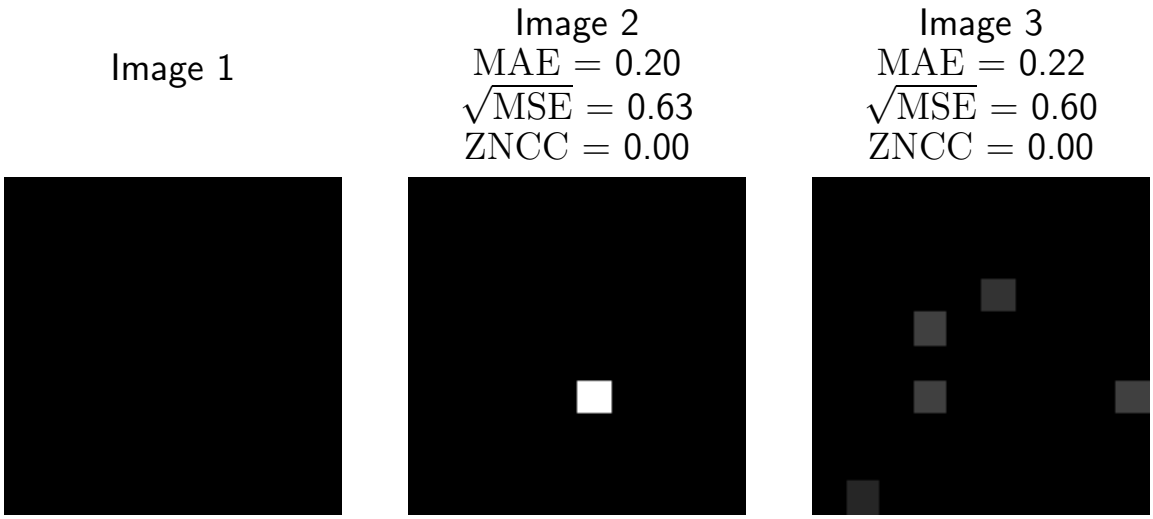


Figure 2.11: *Three images with same size of 10 by 10 (width and height) pixels. Note that ZNCC is multiplied by 100 which the value is displayed in percentage.*

Another example is shown in Figure 2.12. Image 2 is the mean image of Image 1 and Image 3 is computed by subtracting Image 1 by the its mean value. For both metrics (MAE and RMSE), they believe that Image 2 is more similar to Image 1. However, Image 3 appears to be more similar to Image 1 visually which is agreed by ZNCC (100% for Image 1 and 3 vs 0% for Image 1 and 2). From those examples we can conclude that a single metric is generally not very reliable and care must be given to the choice of metrics. Also, visual comparison or assessment can be helpful but expert knowledge is required if complex data is involved (e.g. medical images).

2.5 Survey of optimisation algorithms

Optimisation is a process of finding the best solution or solutions of a given problem. An optimisation algorithm is a set of rules (or instructions) that describes the optimisation process. Optimisation algorithms are often used in solving complex problems that require repetitive

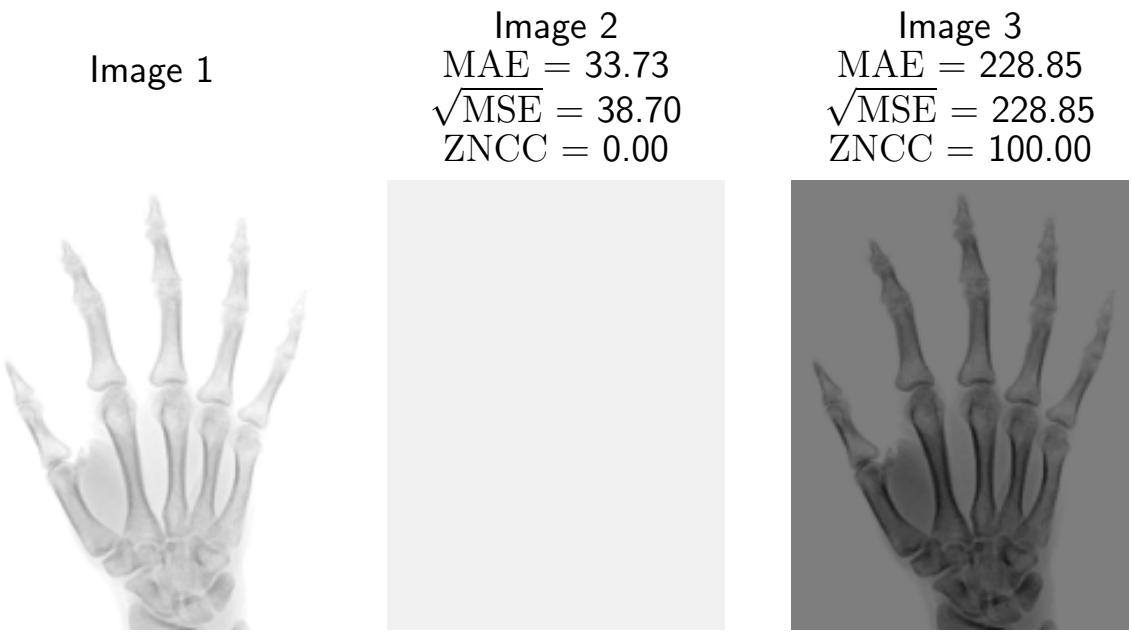


Figure 2.12: *Three images with same size. Image 2 is the mean image of Image 1 and Image 3 is computed by subtracting Image 1 by its mean value. Note that ZNCC is multiplied by 100 which the value is displayed in percentage.*

computation. The main component in an optimisation algorithm is the objective function which quantifies the quality of solutions. The objective function is constructed using one or multiple metrics (see Section 2.4). The goal of optimisation algorithms is to optimise the objective function which will gives the best solution.

2.5.1 Pure random search

The idea of pure random search (PRS) is simple: i) generate *N* random sample points uniformly distributed within given boundaries, and ii) select the solution that produce the best metric value. The best solution must have the lowest metric value (Figure 2.13) or highest metric value if similarity metrics are used. PRS is simple to use and fast to run, and it is often used as a baseline to compare with other methods.

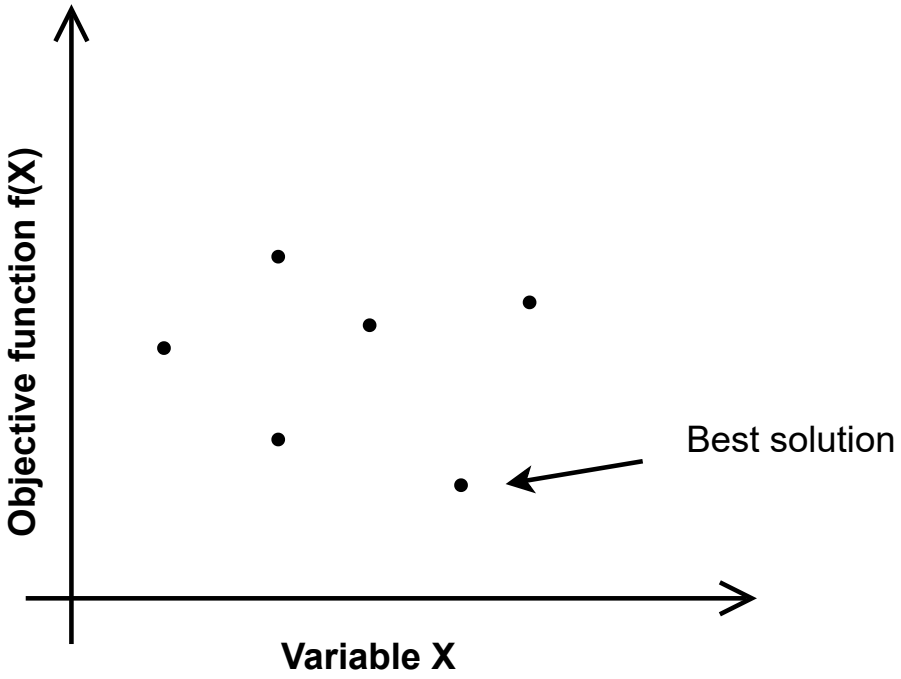


Figure 2.13: *A set of solutions in a minimisation problem where the best solution is selected based on lowest metric values when PRS is used.*

2.5.2 Simulated annealing

An unique current solution is used at any one time in optimisation algorithms such as Downhill Simplex, conjugate gradient (CG) and simulated annealing (SA). However, the introduction of an acceptance probability enables the SA algorithm to accept solutions worse than the current one to escape local optima [95] (see Figure 2.14).

SA is based on annealing in metallurgy where a material is heated and cooled down in a controlled environment. The aim is to reduce defects in the treated material. SA mimics this behaviour with:

- A high temperature is first used. The aim is to widely explore the search space.
- The temperature is progressively reduced to restrict the search space, hence refine the results.

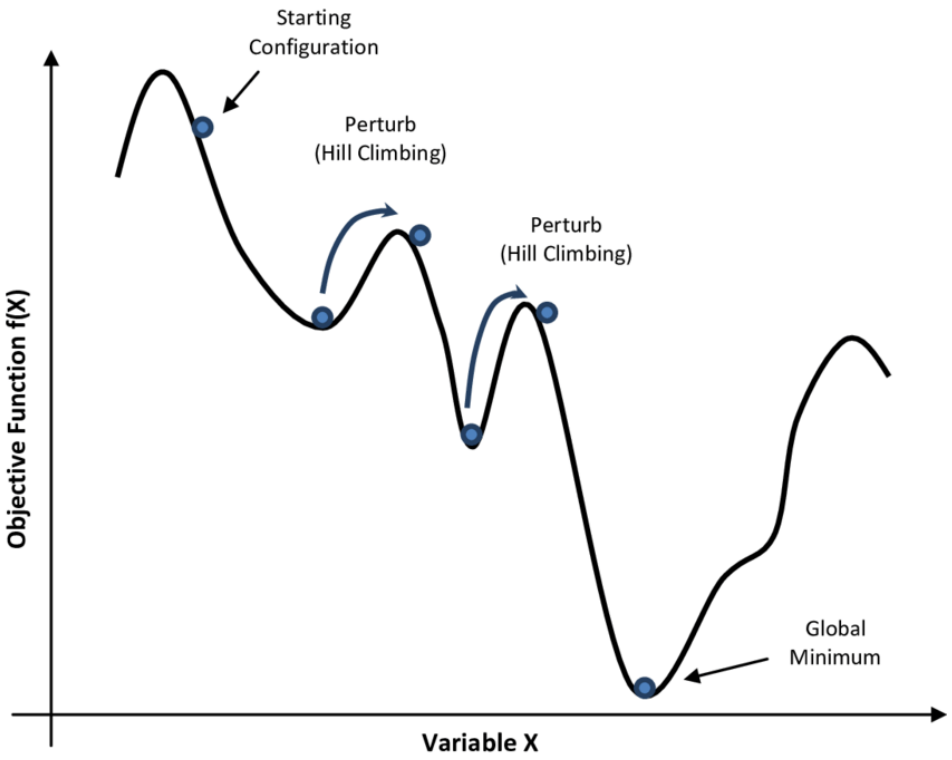


Figure 2.14: *Illustration of the optimisation process in SA. Source: Ghasemalizadeh, Khaleghian and Taheri (2016) [44].*

Solutions could possibly be improved by introducing a restart mechanism where the annealing process is repeated for several times. One disadvantage of SA is that temperature cooling must be slow, i.e. more computational resources are required.

2.5.3 Evolutionary Algorithms

In evolutionary algorithms (EAs) concurrent solutions are used at each iteration of the algorithm [8]. An EA relies on the theory of biological evolution proposed by Charles Darwin [29]. A flowchart of how EA works in the 3D/2D registrations is presented in Figure 2.15. A solution is called 'individual' that is defined by a sequence of genes which is a set of transform parameters in our case. The performance of each individual is measured using a fitness function (i.e. the objective function to be optimised). The set of individuals is called 'population'. Genetic operators (usually selection, crossover, mutation and elitism) are repetitively used on the individuals of the population. Individuals with best fitness values are

given a higher probability to reproduce. At each iteration, a new population of offspring is generated from the current population (the parents) where least-fit individuals are removed. At the end of the algorithm, the best individual is obtained. A restart mechanism that repeats the whole process several times can also be used to increase the chance of getting out of the local optima. The main advantages of EA are no gradient information is needed and the ability to approach global optima.

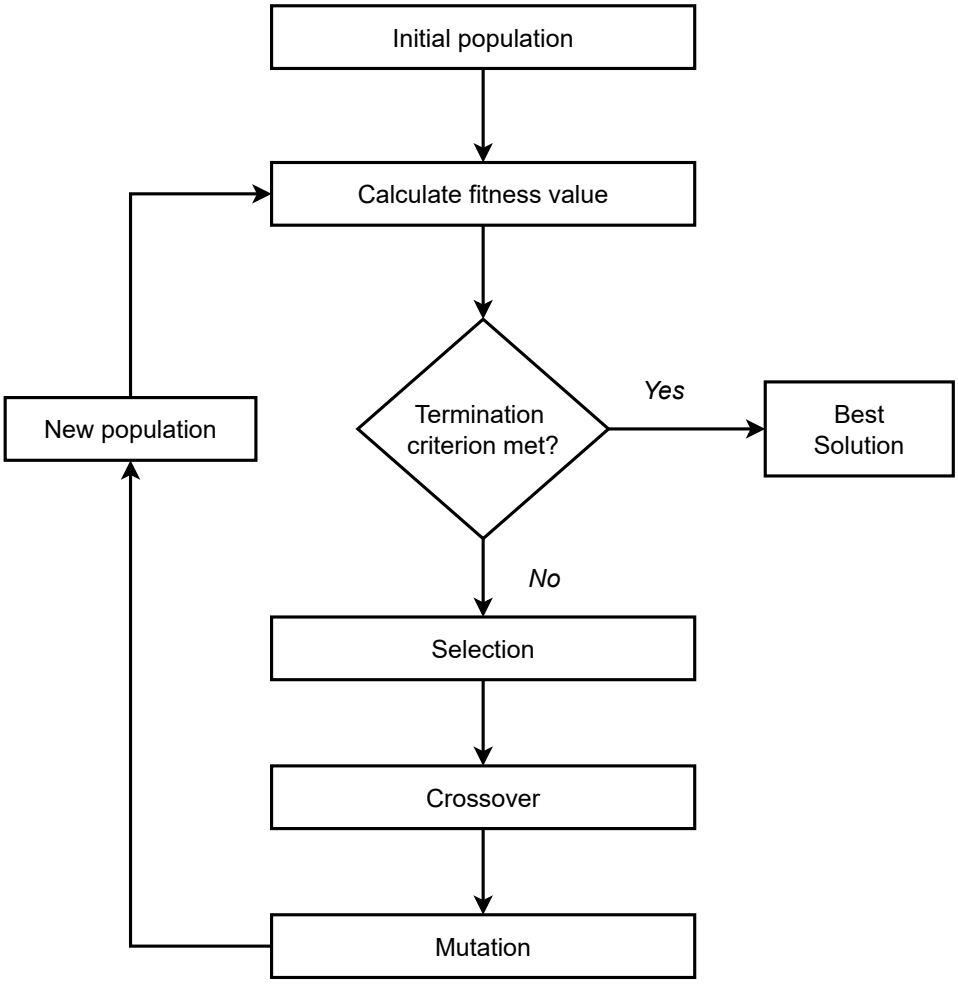


Figure 2.15: *Illustration of the EA algorithm.*

2.5.4 Fly algorithm

The Fly algorithm [87] is an evolutionary algorithm that is based on the Parisian approach. All individuals in the population collaborate toward a common goal rather than compete [27]. It was initially developed for real-time stereo vision in robotics. The Fly algorithm is implemented

as any other EA (i.e. with all the common genetic operators such as selection, crossover and mutation) [49]. It evolves a population of individuals called *flies*, which correspond to 3D points in the problem space. Each fly is projected in the object space. The type of projection is problem-dependent. The fitness function is computed from those projections. In the Parisian approach there are two levels of fitness function:

1. The *local fitness function* evaluates the performance of a given fly. It is used during the selection process. For a fly, improving its local fitness means increasing its chances of survival and being selected to reproduce.

2. The *global fitness function* assesses the performance of the whole population. Improving the global fitness is the goal of the whole population which produces the global optima.

The local fitness can be implemented as a marginal fitness (F_m) using the global fitness with the leave-one-out cross-validation principle (MAE is used in the *global fitness function*):

$$F_m(i) = \text{MAE}(\mathbf{Y}, \hat{\mathbf{Y}} \setminus \{i\}) - \text{MAE}(\mathbf{Y}, \hat{\mathbf{Y}}) \tag{2.21}$$

where where $\hat{\mathbf{Y}}$ is the projections estimated from the whole population (e.g. it is the sinogram simulated from the flies in the X-ray microtomography (micro-CT) registration), $\hat{\mathbf{Y}} \setminus \{i\}$ is the projections estimated from the population without Fly i. The idea behind the leave-one-out cross-validation is to assess the error metric twice: once with Fly i in the population, and once without it. By comparing the two values (the subtraction in Eq. 2.21) we can determine if having Fly i is beneficial or not for the population:

- If F_m is positive, the error is smaller when the fly is included: the fly has a positive impact on the population's performance. It is a good fly which should be a good candidate for reproduction.

- If F_m is negative, the error is larger when the fly is included: the fly has a negative impact on the population's performance. It is a bad fly which should be eliminated.

F_m is therefore a measurement that is maximised by the algorithm whereas the *global fitness* $\text{MAE}(\mathbf{Y}, \hat{\mathbf{Y}})$ is minimised.

2.5.5 Covariance matrix adaptation evolution strategy

Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [51] is a widely used optimisation algorithm. Other optimisation algorithms could be used for further evaluations of our approaches such as multi-objective optimisation algorithms [151]. CMA-ES is a special evolution strategy with adaption of covariance matrix. It is used to solve complex problems that require derivative-free optimisation. An evolution strategy is inspired by biological evolution. The idea is: initialising individuals (a set of solutions), recombination and mutation is used to create new individuals, best individuals are then selected based on their fitness value to become the parents of next generation of individuals. This process is repeated until satisfactory results (set termination criterion) are found (illustration is shown in Figure 2.16). In Evolution Strategies (ES), new individuals are created by sampling from the probability distribution. In CMA-ES, however, sampling is achieved through the use of a covariance matrix of the distribution. This gives CMA-ES great advantages in ill-conditioned problems where small changes of input variables result in large change of output. CMA-ES is generally not suitable for large-scale optimisation problems due to high space and time complexity [66].

2.5.6 Non-dominated Sorting Genetic Algorithm II

Non-dominated Sorting Genetic Algorithm II (NSGA-II) is a multi-objective EA [32]. Unlike algorithms that are introduced in previous sections (such as SA, EA, CMA-ES, etc.), NSGA-II optimises multiple objectives at the same time. Initially, a random population is created. Another offspring population is created by using binary tournament selection, recombination and mutation operators. These two population sets are formed as the initial population which is then sorted based on non-domination. Each individual is assigned a rank value and individuals with low rank values are more likely to be selected. Then the crowded-comparison operator is used to select a list of individuals. Crowded-comparison is the process where the crowding distance between individuals of the population is calculated. The crowding distance is the

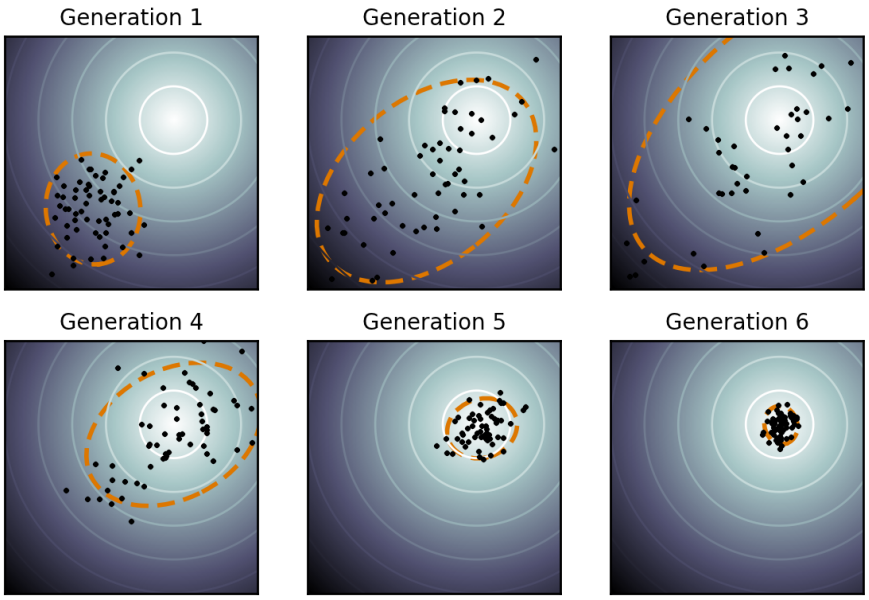


Figure 2.16: *Illustration of CMA-ES on a two-dimensional problem. Points represent individuals in the population and dotted spheres represent the distribution of the population. Source: Wikipedia [169].*

average distance between an individual and its neighbours. Individuals with extreme objective values are assigned distance values of infinity. In this way the diversity of the population is maintained. Finally, a new population is formed with the same number of individuals as the initial random population. The above procedure is repeated until optimal solutions are found.

Usually, there are more than one solution and those solutions are termed as pareto front (PF) or Pareto-optimal set. The PF is a list of non-dominate solutions where every solution is as good as others. To understand the concept we can start with a simple bi-objective optimisation problem. ZNCC and MAE are used to construct the objective function. The goal of the optimisation process is to find a solution with high value of ZNCC and low value of MAE. We obtained a set of 10 solutions (Figure 2.17). Clearly, there is a PF set containing 6 solutions that are represented by black squares. The rest 4 points (red triangles) are excluded because they can always improve on both metrics, i.e. there is always another solution with lower MAE and higher ZNCC. However, solutions in the PF can not be improved further.

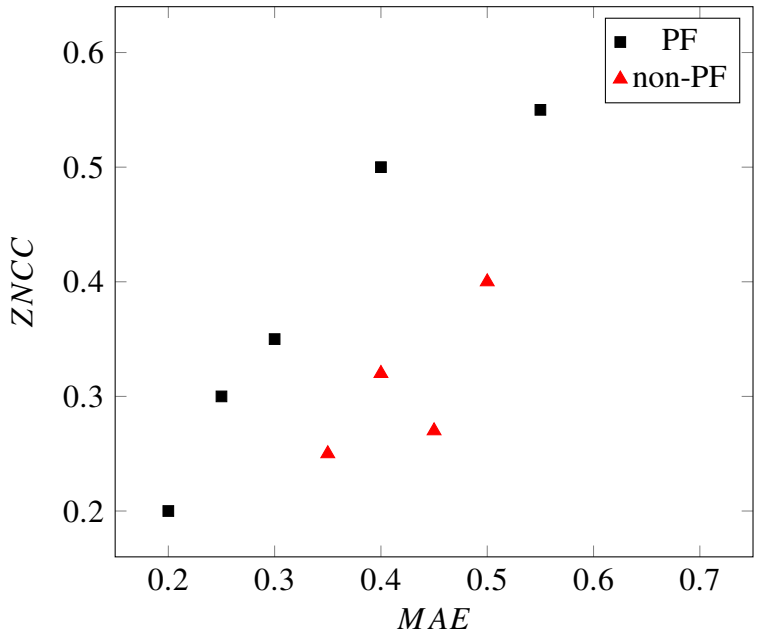


Figure 2.17: *An example set of 10 solutions using NSGA-II with two metrics. The PF contains 6 solutions (black square) and the rest 4 solutions are not in the PF (red triangle).*

2.6 X-ray simulation

X-ray simulation is extensively studied in physics; with application in medicine and material science. Physically-based simulation codes are available [2], [9], [13], [40]. The most popular technique relies on the Monte Carlo (MC) method. It is often used in dosimetry for radiotherapy due to its high level of accuracy. MC-based simulation codes implement probabilistic X-ray interaction models for the transport of photons in matter. X-photons cross matter. During their path into the object, they can interact with matter. Photons can reach the detector without any interaction (see '1' in Figure 2.18). They can be absorbed, in which case they do not reach the detector and do not contribute to the X-ray projection (see '2'). Scattered photons decrease the image quality with noise and blur (see '3'). Photons can also be scattered then absorbed (see '4'), or even be scattered multiple times. At each iteration of the MC algorithm, interaction events may occur between each photon and the material it is crossing based on these probabilistic models. For some applications this approach leads to excessive computation time due to the stochastic nature of MC methods. Days if not weeks may be required to simulate a single X-ray projection with an acceptable signal-to-noise ratio (SNR). For transmission

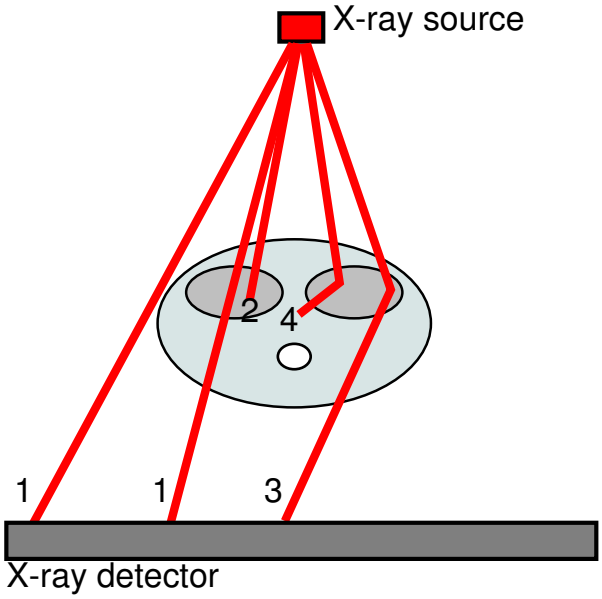


Figure 2.18: *X-photons/matter interactions.*

imaging (inc. radiography, CT, cone beam computed tomography (CBCT), and fluoroscopy), deterministic calculation using the Beer-Lambert law (also known as attenuation law) might be considered as a sufficient description for the expected value of the number of photons that did not interact with the objects:

$$N_{out}(E) = N_{in}(E) \exp\left(-\sum_{i} \mu_i\left(E, \rho, Z\right) L_p(i)\right) \tag{2.22}$$

with $N_{in}(E)$ the number of incident photons at energy E, $N_{out}(E)$ the number of transmitted photons of energy E, μ_i the linear attenuation coefficient (in cm^{-1}) of the i^{th} object and $L_p(i)$ the path length of the ray in the i^{th} object. μ_i depends on E the energy of incident photons, ρ the material density of the object, and Z the atomic number of the object material.

When material composition of scanned object is homogeneous, the total energy $E_{out}(x, y)$ that is received at the detector pixel position (x, y) can be derived:

$$E_{out}(x, y) = E \times N_{in} \times \exp\left(-\sum_{i=0}^{i<objs} \mu(i, E) L_p(i, x, y)\right) \tag{2.23}$$

Where $objs$ is the total number of scanned objects, $L_p(i, x, y)$ is the path length of the ray in i^{th} objects where the ray is formed by connecting the point source with the position of pixel (x, y) on the detector, $\mu(i, E)$ is the linear attenuation coefficient at energy E for the i^{th} object.

In a polychromatic case where there are different energies M in the incident beam spectrum, the total energy $E_{out}(x, y)$ can be defined as:

$$E_{out}(x, y) = \sum_{j=0}^{j<M} E_j \times N_{in}(E_j) \times \exp\left(-\sum_{i=0}^{i<objs} \mu(i, E_j) L_p(i, x, y)\right) \tag{2.24}$$

When a set of points is used to construct the X-ray source (instead of a single infinitely small point), the total energy $E_{out}(x, y)$ becomes:

$$E_{out}(x, y) = \sum_{k=0}^{k<p} \sum_{j=0}^{j<M} E_j \times N_{in}(E_j) \times \exp\left(-\sum_{i=0}^{i<objs} \mu(i, E_j) L_p(i, k, x, y)\right) \tag{2.25}$$

In the X-ray simulation three inputs need to be defined to generate X-ray images:

X-ray source is defined by by its shape, position, orientation, and incident photon beam. The shape can be a single infinitely small point, a set of points, or others. The incident photon beam is determined by its spectrum (i.e. a list of photons) where it can either be *monochromatic* (same energy) or *polychromatic* (different energies).

X-ray detector is defined by the number of pixels along the horizontal and vertical axes, its position and orientation. Usually, the number of pixels along both axes are the same but it is not always the case. The direction of X-rays are perpendicular to the detector in most cases although it doesn't have to.

Scanned objects are defined by its shape that is represented by polygon meshes and their material properties in term of density and chemical composition. Hounsfield values can be

assigned to each polygon mesh where those values can be converted into density and chemical composition [133]. Material properties of objects are provided by the XCOM database from the National Institute of Standards and Technology (NIST) [11].

In this case, methods based on the ray-tracing principle are often used as a fast alternative to MC methods [40]. To speed-up computations, GPU can be used, but often focusing on radiotherapy and voxelised data [12], [64]. For this 3D/2D registration problem, we rely on the simulation of X-ray images from polygon meshes. We make use of gVirtualXRay, a cross-platform Open-Source library [156], [161]. It implements a multi-pass renderer to solve one of three equations (2.23, 2.24 and 2.25). It is written in C++ using OpenGL and the OpenGL Shading Language (GLSL) to provide real-time performance. Wrappers to other popular languages are also provided. Here we use the Python3 wrapper. Technical details can be found in [157], [158], [161]. gVirtualXRay has been used in various medical applications where speed and accuracy are both requirements, including radiograph generation from dynamic polygon meshes (to simulate patient posing and respiration) [144], [158] and the simulation of CT data acquisition and reconstruction including beam hardening or motion artefacts due to the respiration [160].

gVirtualXRay has been validated against results that are obtained MC simulation toolkit developed by the European Organization for Nuclear Research (CERN): Geant4 [3]. Simulated X-ray images are almost identical where NCC between two images is 99.747%. Other validation tests have been conducted [157], [161]. More recently, we submitted a journal article where the CT scan of the Lungman phantom is segmented to extract polygon meshes. A simulated radiograph is then registered on a real digital radiograph of the same phantom. The ZNCC between two radiographs with anteroposterior (AP) view is 99.36%.

2.7 Discussion and Conclusion

We have discussed the general concept of registration and different strategies that are used to establish dimensional correspondences in 3D/2D registration methods. Different types of 3D/2D registration methods were presented including *intensity-based* registration methods that

utilise the projection strategy. The main drawback of *intensity-based* registration methods is insufficient number of projection views would affect the registration performance. However, improvements on registration accuracy became marginal when number of views exceeds three.

We have also discussed how deep learning techniques can be applied to solve registration problems, which refers to *DL-based* 3D/2D registration methods. Deep learning has been successfully deployed to solve many complex problems. However, there are some disadvantages of *DL-based* methods such as the requirements of large amount of training data and computational resources. For real-life complex problems, a lot of transformation parameters are required. This means the amount of annotated training data is vast even if synthetic images are used, e.g. Miao *et al.* used 25,000 synthetic image for 6 transformation parameters (see Section 2.3). The X-ray source and detector are fixed. If one of them moves, additional training data will be required. Also, registration methods that rely on extracting features from original images might not perform well when the position of X-ray source and detector is unknown. The object in the image might become too small to work with which might lead to bad registration performance.

Some popular similarity and dissimilarity metrics that are used to construct the objective function are presented. Traditional optimisation algorithms such as PRS, more advanced single-objective optimisation algorithms such as fly algorithm and CMA-ES and bi- or multi-objective optimisation algorithms such as NSGA-II that aim to optimise the objective function were also presented. Single-objective optimisation algorithms are often easier to use comparing with bi- or multi-objective optimisation algorithms especially in choosing the best solution from the PF. It is necessary to consider techniques that can help with solution selection. In the next chapter we will introduce our 3D/2D registration method which heavily relies on some of optimisation algorithms and objective functions that are presented in this chapter.

Chapter 3

Registration of synthetic and clinical X-ray radiographs

3.1 Introduction

In the previous chapter we have discussed the concept of 3D/2D registration and different types of 3D/2D registration methods. In a 3D/2D registration problem, one dataset has to be transformed into the same coordinate system to match another. It is achieved through establishing the dimensional correspondence between 3D and 2D data. There are three strategies can be applied. The projection strategy transforms 3D data onto 2D plane where 3D/2D registration becomes multiple 2D/2D registrations. The back-projection strategy and the reconstruction strategy transform 2D data into 3D space where 3D/2D registration becomes multiple 3D/3D registrations. Some registration methods based on deep learning has been discussed. We also discussed some similarity and dissimilarity measurements for matching the DRR that generated from a 3D dataset with an X-ray image. Registration results were obtained by minimising the dissimilarity metric or maximising the similarity metric between the DRR and the X-ray image. Some popular optimisation algorithms were presented.

Computational methods of disease tracking and progression prediction based on the analysis of medical imagery is receiving heightened attention in recent years. Chronic diseases of the human musculoskeletal system caused by autoimmune processes lead to progressive, irreversible anatomical changes over time. In the case of rheumatoid arthritis (RA), a chronic inflammatory

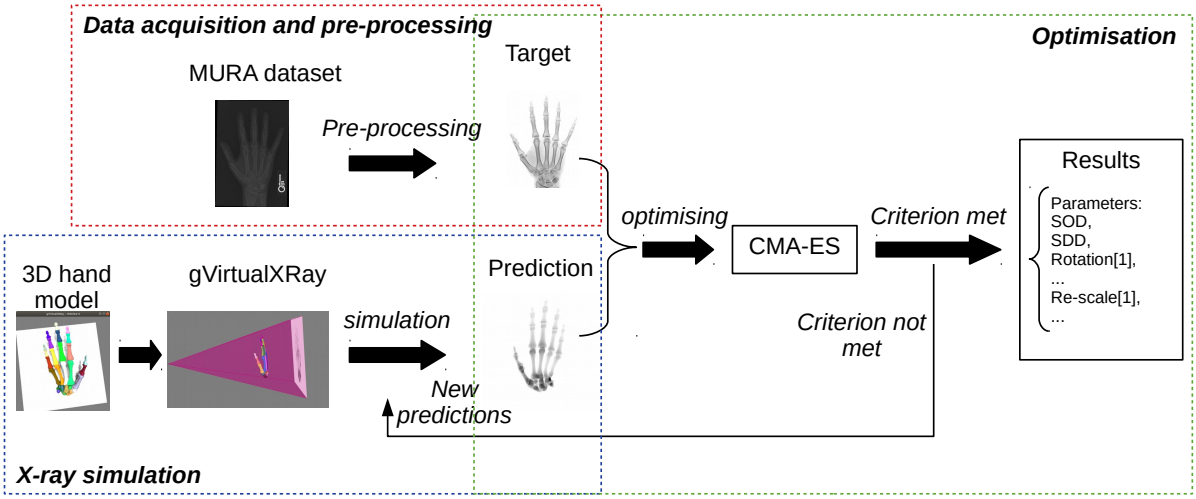


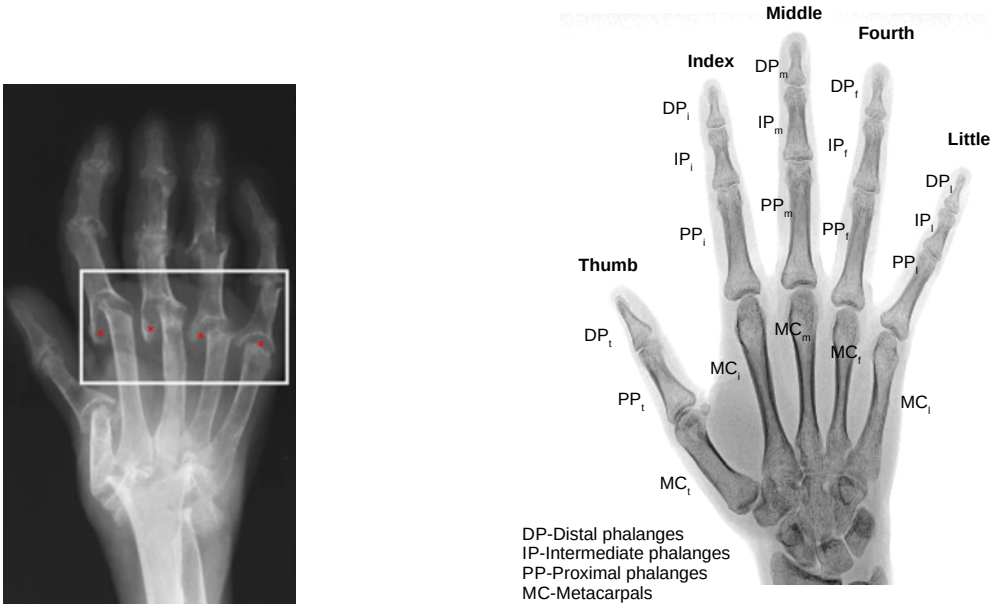
Figure 3.1: *Registration pipeline based on X-ray simulation and CMA-ES.*

disorder with largely unknown pathogenesis, patients often present to the clinician with swelling of the hands. If left untreated, the disease progresses in distinct stages, from joint pain, swelling, stiffness to cartilage loss, bone erosion, deformities and total loss of joint function [106].

Plain radiographic imaging (X-rays) of the hands is done routinely for diagnostic and tracking purposes, as routine care for RA patients. Since hand radiographs are relatively inexpensive and low-risk, they provide clinicians with baselines, and visible changes over time. The rate of disease progression is modulated by treatment and lifestyle choices, but distinct deformations have been documented [108], [131]. Typical deformities include boutonnière, swan-neck, hitchhiker's thumb and claw toe. Other, less obvious changes, include bone erosions, induced by the inflammation of the synovial membrane, as shown in Figure 3.2a.

Radiographs are projections of 3D structures, hence much information is lost. The anatomy of the hand varies among different individuals, e.g. the ratio of the lengths of the long bones is not always consistent. Using this observation and the ability to speedily create DRRs, we propose a method to register a 3D mesh model of a hand to PA view hand radiograph.

Anomaly detection in hand radiographs is important for disease staging and monitoring. Our registration method is a pre-processing step for algorithms that modify the mesh model using domain-specific knowledge to better track disease-induced changes without expensive



(a) **(b)**

Figure 3.2: *a) Erosions induced by RA inflammatory processes visible around the red asterisks. Image source [131]. b) Naming of hand fingers and bones. Source: The MURA dataset.*

volumetric scans that clinicians may not be equipped with or are cost-prohibitive. Time-series, patient-specific information regarding the progression of a disease is critical for treatment planning and drug effectiveness monitoring.

In this chapter we present a 3D/2D registration method where the 3D model is transformed so that their corresponding X-ray projection(s) finely match an input image. It is based on the projection strategy (Figure 2.2). Our framework heavily relies on optimisation techniques for the tuning of transformation parameters and on fast X-ray simulation on GPU for the generation of X-ray projections. In most 3D/2D registration problems, the source corresponds to a 3D volume (e.g. voxels from a CT or MRI dataset) and the target is a 2D image (e.g. an X-ray image) (Table 2.1).

In our application the source is a scenegraph containing multiple 3D triangular models, and the target corresponds to an X-ray radiograph or a sinogram (a collection of projections at different angles). Every 3D model can be deformed in the nonrigid framework. Our hypothesis is that fast X-ray simulation from polygons on the GPU can be embedded in objective functions to

transform 3D surface models using black-box optimisation. We firstly perform the registration of a 3D hand model to synthetic (or simulated) radiographs as a proof-of-concept study. Then we perform registrations on clinical radiographs to demonstrate the feasibility of our method in real-life cases.

The rest of the chapter is organised as follows: Section 3.2 describes our method and presents preliminary results on registering synthetic images, Section 3.3 describes changes to our method including the data used and data pre-processing, and presents results on registering real clinical radiographs and Section 3.4 summaries our findings and discuss further research direction.

3.2 Preliminary study using synthetic X-ray images

3.2.1 Introduction

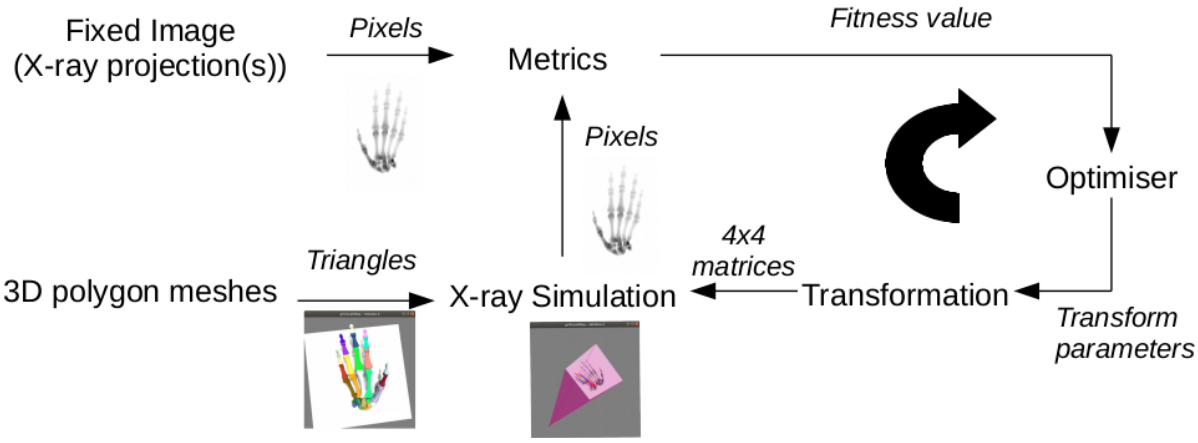


Figure 3.3: *Registration pipeline based on X-ray simulation and black-box optimisation techniques.*

Figure 3.3 provides a summary of our approach. It corresponds to a 3D/2D registration problem where a moving dataset (called source) is deformed to match another one, which is fixed (called target). In most 3D/2D registration problems, the source corresponds to a volume dataset (e.g. from a CT or MR images) and the target is a 2D image. In our application the source is a scenegraph containing 3D triangular models, and the target is a simulated 2D image (X-ray radiograph or sinogram). The 3D models are deformed using rigid body transformations applied on the triangular meshes. The parameters of the transformations are finely tuned to minimise

the difference between the target and the image simulated using the deformed models. This parameter tuning is achieved using black-box optimisation. Note that it is also possible to maximise the correlation (or any other measurement of similarity) between the two images.

The implementation of the registration is written in Python using the *minimize* package provided by SciPy [67] and the implementation of stochastic optimisation methods available on GitHub [89]. We aim to demonstrate that the choice of optimisation methods is critical when developing a registration framework.

We model the registration as a black-box optimisation problem where a simulated image (called $source$) should accurately match a given input image (called $target$). These image correspond to a radiograph for the hand problem. The parameters used to simulate the X-ray projections are finely tuned by an optimisation algorithm using an objective function. In minimisation, the objective function corresponds to a distance metrics between the $source$ and $target$. The optimisation algorithm aims to lower its value as much as possible. In maximisation, similarity or correlation metrics are used instead and the optimisation algorithm aims to increase its value. In black-box optimisation, no assumption is made on the shape of the objective function and no gradient information is used, which is the case in our registration problems.

As the objective of this section is to find suitable optimisation algorithms, the metrics must be robust and easy to implement. There are two popular metrics for comparing true and predicted values: one is MSE (see Section 2.4.6) and another one is MAE. MAE is more favourable than MSE since MSE relies on squared values which penalise outliers (large values). The objective function that we minimise corresponds to the MAE is shown in Equation 2.19. For the purpose of comparison, lower MAE values can typically be interpreted as better predictions: MAE is equal to zero when $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ are strictly identical.

3.2.2 Results

The hand registration problem is a toy problem to assess the feasibility of our framework. The scenegraph of a 3D hand model is used (see Figure 3.4). Some of the nodes, including the root

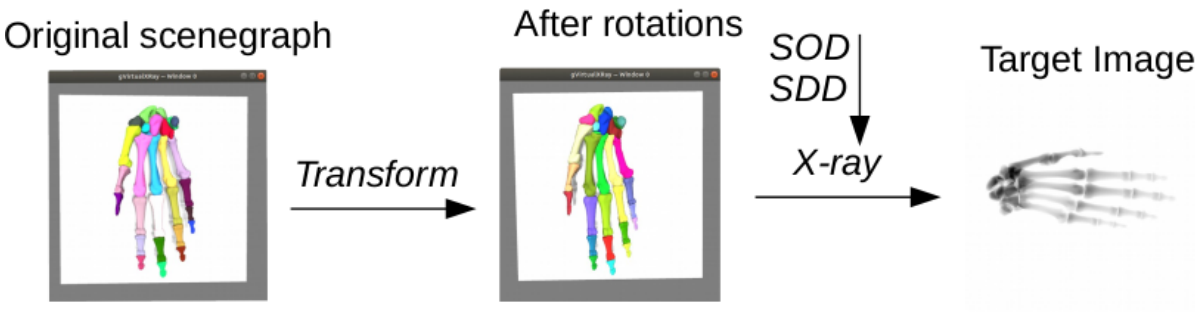


Figure 3.4: *Original and target hand poses for registration.*

node, are arbitrary rotated so that the fingers are now close to each other and the whole hand is tilted. We chose this pose to check the ability of the optimisation algorithms to retrieve subtle changes. The SDD and SOD are set. The corresponding X-ray radiograph is then computed: it is the target image in the registration. The transformations mentioned above, including SDD and SOD, are now 'forgotten'. The aim of the registration is to retrieve all these parameters. For fair comparison, each methods are stopped after the objective function is called 500 times, except for SA where the objective function is called 508 times. In EA, 25 individuals are generated and running for 20 generations.

Figure 3.5 shows the performance of each method performing on the hand registration problem over 15 runs. EA outperforms others with lowest median value of distance. The highest MAE for EA is still lower than the lowest distance for SA, PRS and CG. It also has good stability, which is at an advantage in medical field. Nelder-Mead provides good results but comes with a higher standard deviation, which means it can fail to find the global optimum when it gets stuck in a local optimum. SA performs close to Nelder-Mead but comes with narrow spread. It might be a better alternative for variation-sensitive problems. Both PRS and CG do not provide very good results in this problem.

To further refine our analysis, the visualisation of the error map between target and prediction (median solution) for each method is provided in Figure 3.6. The error maps are normalised between 0 and 1, and visualised using the 'thermal' colour look-up table provided by Fiji [132]. Our aim is to highlight in each figure, independently from one another, where the error is the highest. EA and Nelder-Mead provide the best median results in term of MAE. EA's error

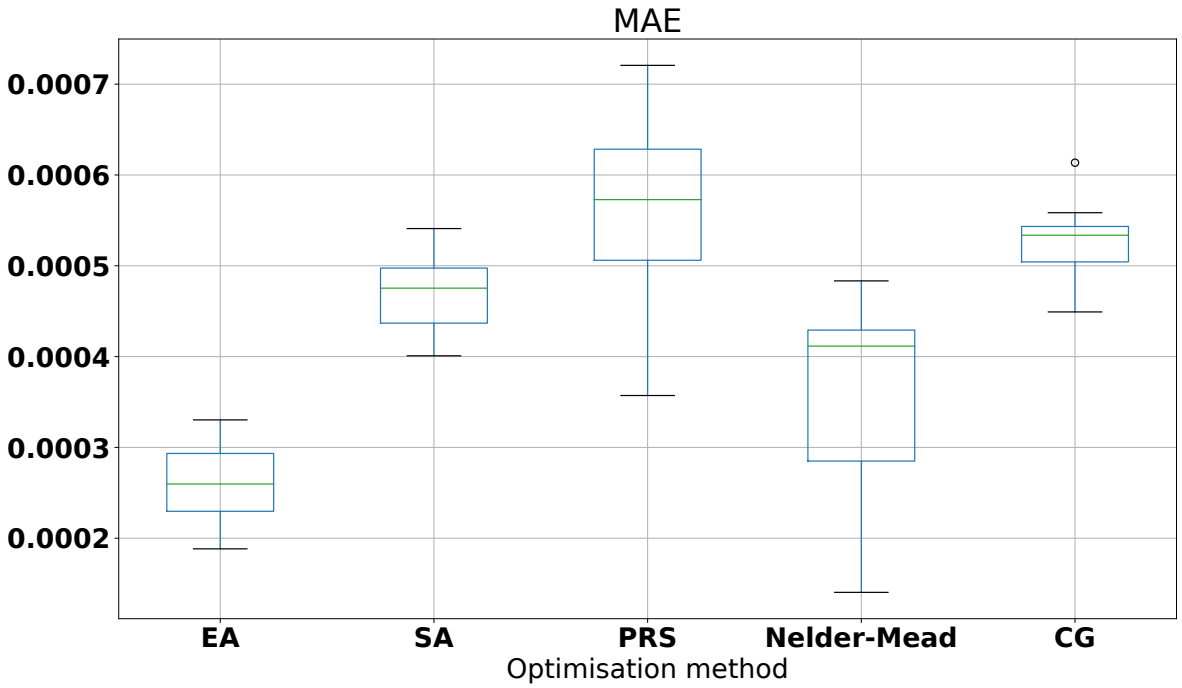


Figure 3.5: *Overall performance over 15 runs of each method on the hand registration problem. Each optimisation algorithm called the objective function about 500 times. A good algorithm to solve this problem proves a low MAE and exhibits a small spread over the 15 runs (in other words it consistently provides results of good quality).*

map shows that the error is concentrated in the thumb. The fingers are now close to each other as in the target image. For Nelder-Mead, the error is concentrated in the middle and ring fingers. SA provides acceptable results. CG and PRS clearly provide the worse results, which are in line with our expectations. Although the value of MAE are similar for CG and PRS, the predicted images varies slightly. This shows the limit of comparison metrics as discussed in Section 2.4.14.

3.3 Registration of clinical radiographs

3.3.1 MURA dataset

The MURA dataset, freely available on GitHub at `https://stanfordmlgroup.github.io/competitions/mura/`, contains 40,561 musculoskeletal radiographs from 14,863 clinical studies of 12,173 different patients [117]. It focus on upper extremities of human body, including elbow, finger, forearm, hand, humerus, shoulders and wrist. Each radiograph is manually labelled as normal or abnormal. 15 radiographs (Figure 3.7) are selected to test the performance of our approach (Section 3.3.4.1), which only require a single radiograph.

Unlike typical medical images such as Digital Imaging and Communications in Medicine (DICOM) [48], radiographs in the MURA dataset are in Portable Network Graphics (PNG) format in which a lot of information is missing (Table 3.1). To adapt to the data to use in our experiments, those radiographs are pre-processed to improve the structural difference between the object and other areas. More details about image pre-processing is presented in Section 3.3.2.

3.3.2 Image pre-processing

Although the MURA dataset was created for the purpose of abnormality detection, it is also useful for testing the performance of our registration framework. We focus on hand radiographs with PA views. We manually pre-processed 15 different radiographs as follows (illustration is displayed in Figure 3.8):

Table 3.1: *Information needed to simulate X-ray radiographs that are present in typical medical imaging file formats such as DICOM but missing in the PNG files from the MURA dataset.*

Properties	PNG	DICOM
X-ray tube voltage	plausible energy beam	known
Quantisation	8-bit	16-bit
SOD	need to estimate	known
SDD	need to estimate	known
Pixel spacing (mm)	need to estimate	known
Object location	need to estimate	known
Object orientation	need to estimate	known

1. each radiograph is cropped so that only the hand part remains,

2. the left or right marker is removed and the area surrounding the hand is "cleaned" (to have same pixel values),

3. the skin around finger is removed as much as possible,

4. the radiographs, which are negative images, are inverted to match the positive images generated by X-ray simulation, and

5. all images are re-scaled to the same size of simulated X-ray images.

Corresponding images are shown in Figure 3.9. In each image, we define the name of each finger (from left to right): thumb, index finger, middle finger, the fourth finger and little finger, as shown in Figure 3.2b.

3.3.3 Optimisation

CMA-ES [51] is a widely used optimisation algorithm which provides a great baseline result. Other optimisation algorithms could be used for further evaluations of our approaches such as multi-objective optimisation algorithms [151]. CMA-ES is a special evolution strategy with adaption of covariance matrix. It is used to solve complex problems that require derivative-free

Table 3.2: *Rotation and re-scaling parameters to be optimised and corresponding ranges.*

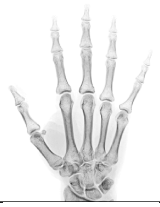
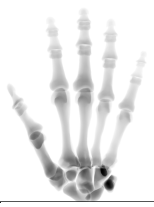
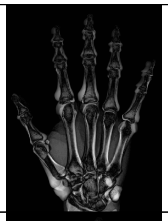
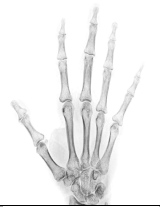
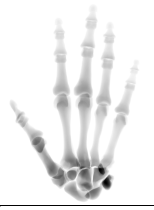
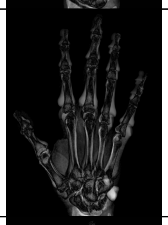
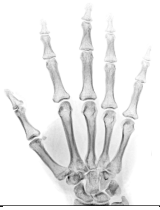
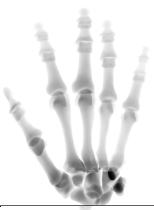
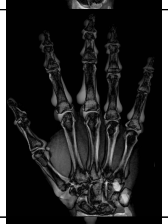
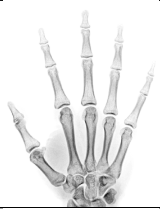
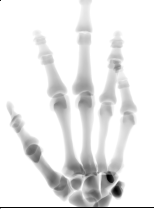
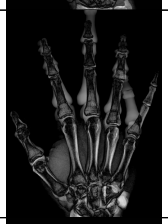
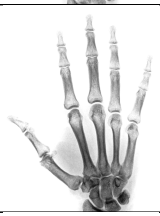
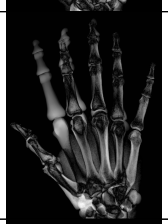
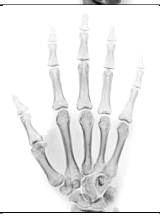
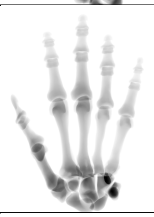
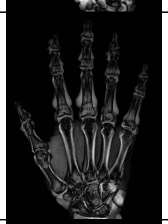
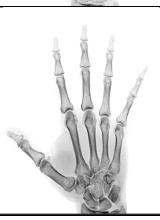
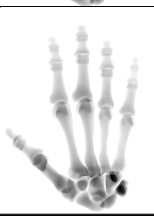
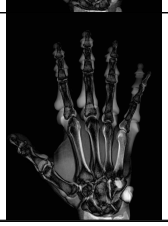
Parameters \ Bones	**Whole hand**	**Thumb**	**Index**	**Middle**	**Fourth**	**Little**
Rotation range (degrees)	[-20, 20]	MC_t: $[-10, 10]$ $[-20, 0]$ PP_t: $[-10, 10]$	PP_i: $[-10, 10]$ $[-20, 0]$ IP_i: $[-20, 0]$ DP_i: $[-20, 0]$	PP_m: $[-10, 10]$ $[-20, 0]$ IP_m: $[-20, 0]$ DP_m: $[-20, 0]$	PP_f: $[-10, 10]$ $[-20, 0]$ IP_f: $[-20, 0]$ DP_f: $[-20, 0]$	PP_l: $[-10, 10]$ $[-20, 0]$ IP_l: $[-20, 0]$ DP_l: $[-20, 0]$
Rescaling ratio	-	$[0.9, 1.1]$	$[0.9, 1.1]$	$[0.9, 1.1]$	$[0.9, 1.1]$	$[0.9, 1.1]$

optimisation. An evolution strategy is inspired by biological evolution. Further details about CMA-ES can be found in 2.5.5. The optimisation process is shown in Figure 3.1 (green box).

Before the optimisation, all images are normalised to have zero-mean and unit-variance. This is to prevent that some features becoming too dominant during optimisation while other features would be less relevant. We use MAE to construct the objective function, i.e. as the fitness function to be minimised by CMA-ES. MAE is the sum of absolute errors between samples and then divided by total number of samples (Eq. 2.19). The best number, that can be achieved is zero. Typically, lower MAE value indicates better optimising result.

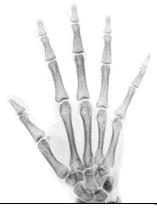
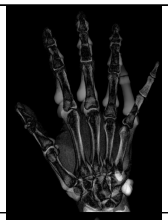
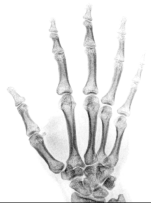
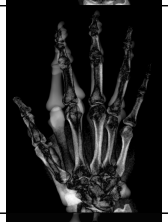
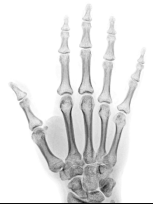
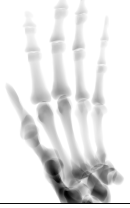
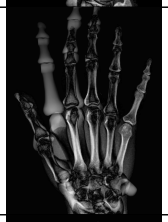
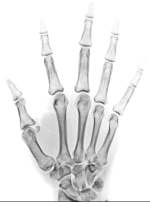
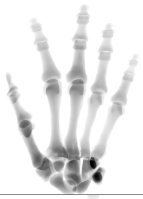
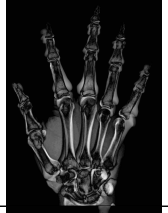
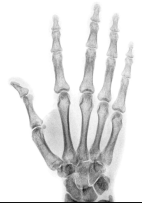
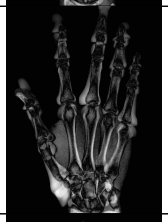
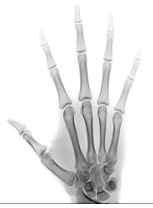
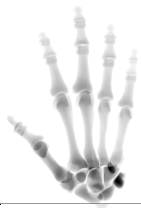
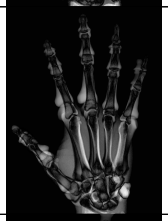
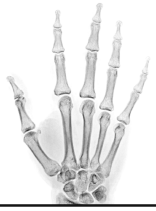
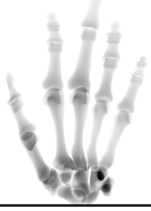
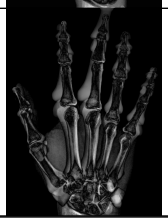
Our registration problem is considered to be complex both in terms of number of parameters and the corresponding data range. There are 38 parameters that need to be optimised including 2 distance parameters: SOD and SDD. SOD is a ratio of SDD with a value between 0.7 and 0.95. By using ratios, we make sure that the distance between the source and the object, is always less than the distance between the source and the detector. SDD ranges between 10 and 1000 centimetres. There are 22 rotating angles and 14 rescaling factors, which are shown in Table 3.2. The rotation range is determined based on the modelling of rotations of the real hand except the whole hand, which is determined by a priori knowledge of the PA pose but adding some degrees of complexity.

Table 3.3: *Registration results for 15 different target images along with corresponding metric values.*

Image number & Metrics	Target	Prediction	Error map	Number of objective function calls
1 & MAE=0.3937 ZNCC=0.7060				2050
2 & MAE=0.3497 ZNCC=0.7378				1964
3 & MAE=0.4060 ZNCC=0.7264				1934
4 & MAE=0.3953 ZNCC=0.7051				2056
5 & MAE=0.4398 ZNCC=0.6603				2050
6 & MAE=0.3947 ZNCC=0.7179				2056
7 & MAE=0.4029 ZNCC=0.7012				2055

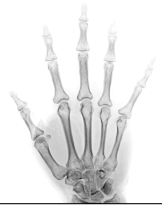
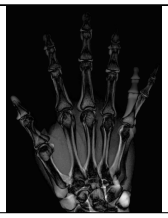
Continued on next page

Table 3.3 – continued from previous page

Image number & Metrics	Target	Prediction	Error map	Number of objective function calls
8 & MAE=0.3765 ZNCC=0.7202				1886
9 & MAE=0.4356 ZNCC=0.6130				2119
10 & MAE=0.4790 ZNCC=0.6503				2308
11 & MAE=0.4167 ZNCC=0.7094				2186
12 & MAE=0.4073 ZNCC=0.7156				1817
13 & MAE=0.3903 ZNCC=0.7218				2016
14 & MAE=0.4464 ZNCC=0.6535				1898

Continued on next page

Table 3.3 – continued from previous page

Image number & Metrics	Target	Prediction	Error map	Number of objective function calls
15 & MAE=0.4206 ZNCC=0.7409				2027

3.3.4 Results

There are two ways to assess the effectiveness of our method in solving the registration problem. In any case, several runs must be performed to gather statistically meaningful data. In Section 3.3.4.1, we selected 15 different radiographs and tested our method once on each of the radiographs. The emphasis is on **data and simulation variability: For different input images, does the algorithm always provide outputs of similar quality?** In Section 3.3.4.2, we selected the images of the worse, median and best registrations of Section 3.3.4.1. The registration is then repeated 15 times for these three images. The emphasis is on **optimisation algorithm variability: For a given input image, does the algorithm always provide a similar output?** We also aim to determine if some images harder to register than others.

3.3.4.1 Data and simulation variability

Here, we aim to determine if the algorithm always provide outputs of similar quality on different input images. 15 registrations using 15 different real X-ray images were performed, i.e. one registration per image, due to computational demand (about 4 hours per registration on a single Intel Core i5-8400 (2.80GHz) central processing unit (CPU) and a single NVIDIA GeForce GTX 1070 Ti GPU).

The 15 pre-processed images that we used are shown in Figure 3.9. MAE is used to compare target and predicted images during registration because it is relatively faster to compute. ZNCC is used for visual analysis of the predicted images after registration. It is a measurement of

similarity between two images. Since it is hard to interpret the value of MAE, ZNCC is very helpful to analyse the performance of the registrations.

To compute ZNCC (see Eq. 2.2), the target and predicted images are normalised first, which is subtracting all pixels by the mean value and divided by standard deviation. Normalised target and predicted images are then multiplied. Finally, all values are added and divided by total number of pixels. ZNCC primarily concentrates on template matching and completely different images might have very high scores.

Table 3.3 lists results from 15 registrations. By looking at predictions and associated error maps, there are 6 registrations that successfully recovered all 5 fingers, where ZNCC is all above 0.7. There are 6 registrations that successfully recovered 4 fingers, where ZNCC is all above 0.7 except image 14. There are 3 registrations that recovered 1 finger, where ZNCC is all below 0.7. It is clear from Table 3.3 that Images 5, 9, 10 and 14 are visually worse than the other images. This trend is not necessarily visible in Figure 3.10a (bar chart of the MAE for each registration). However, Figure 3.10b clearly show two groups: Images 5, 9, 10 and 14 exhibit a significantly lower ZNCC than the other images.

Results demonstrate that our approaches perform well. However, there are some problems that need to be addressed in future researches:

1. Some fingers are not within images. For example, Registration 4, 12 and 14 have the middle finger extended outside the image space. Registration 10 has middle finger and fourth finger extended outside the image space.

2. In 3 registrations (5, 9 and 10) only the thumb is recovered, the middle finger is matched with the target's index finger, the fourth finger is matched with the target's middle finger, and the little finger is matched with the target's fourth finger.

3. A finger is overlapped with another finger. In Registrations 4 and 8, the little finger is overlapped with the fourth finger. In Registration 9, the index finger is overlapped with middle finger.

3.3.4.2 Optimisation algorithm variability

Here, we aim to determine if the algorithm always provides outputs of similar quality on the same input image. We selected the best, median and worse results from the previous subsection, i.e. Images 2, 3, and 10. We perform another 14 registrations on each image and included their previous results (i.e. a total of 15 results per image) to test the variability of the algorithm, CMA-ES, when the input data is the same. Then we compute the mean and standard deviations (STDEVs) of MAE, ZNCC and number of calls to objective function over the 15 runs. The data is summarised in Table 3.4. It shows that CMA-ES provides registrations of consistent quality, both in terms of MAE and ZNCC, for Images 2 and 3 (low standard deviations). However, the standard deviations are much higher for Image 10 (the worse registration of Section 3.3.4.1). The MAE is higher than for the other two images, and the ZNCC lower. It indicates that, somehow, Image 10 is a lot harder to register than Images 2 and 3. Also, differences in those images such as different position of hands might lead to different error being computed, i.e. the position of hands is fixed which only rotation, translation and rescaling is allowed, so any exiting misalignment of hands (e.g. bones on the palm) can lead to greater error. The scatter plot in Figure 3.11 shows the MAE (circles) and ZNCC (triangles) plotted as a function of the number of generations. Green and purple marks are aligned and form horizontal lines: CMA-ES produces consistent registrations for Images 2 and 3. Blue marks are scattered over the plot: CMA-ES does not produce consistent registrations for Image 10.

CMA-ES can produce registrations of good quality consistently for some images. For other images, CMA-ES may fail. Further research is needed to comprehend what makes Image 10 hard to register compared to Images 2 and 3 as the three images are visually similar.

Table 3.4: *Results for 15 registrations on Images 2, 3 and 10.*

Image number	MAE (mean ± STDEV)	ZNCC (mean ± STDEV)	Number of objective function calls (mean ± STDEV)
2 (best run in Table 3.3)	0.3460 ± 0.0055	0.7426 ± 0.0074	2105 ± 189
3 (median run in Table 3.3)	0.4076 ± 0.0019	0.7227 ± 0.0035	1954 ± 132
10 (worse run in Table 3.3)	0.4701 ± 0.0218	0.6597 ± 0.0184	2137 ± 162

3.4 Discussion and conclusion

Our research hypothesis was that 3D/2D registration of 3D triangular meshes onto a 2D image can be performed using optimisation and fast X-ray simulation on GPU. We have implemented a registration framework in Python using open-source technologies, namely SciPy (conjugate gradient and Nelder-Mead) and our own implementation of pure random search, simulated annealing, evolutionary algorithm and Fly algorithm for the optimisation, and gVirtualXRay for the X-ray simulation on GPU. The framework has been successfully tested using synthetic radiographs to perform hand registration for a rheumatoid arthritis application.

To evaluate our pipeline, each optimisation has been repeated 15 times to gather statistically meaningful results. The outcome shows that:

- Care must be given when choosing an optimisation algorithm as some traditional and popular techniques may lead to unreliable results due to local optima;
- Stochastic algorithms such as simulated annealing and evolutionary algorithm can outperform conjugate gradient and Nelder-Mead in term of accuracy and/or reliability;
- X-ray simulations can be repeatedly computed in objective functions thanks to the computational power provided by modern GPUs.

These preliminary results demonstrated the validity of our approach as the registrations were performed successfully.

The second part of this chapter concentrated on evaluating the performance of our method using 15 real clinical X-ray images. We used CMA-ES, a popular evolutionary algorithm for non-linear or non-convex continuous optimisation problems. Results showed that registrations are not always with the same quality but CMA-ES produces similar results with low variability during 15 runs on the same image. This demonstrate that the stochastic nature of CMA-ES is not a concern in our case. It also indicates that some images are harder than others to register.

In the future, we will use this work as an initialisation step in a pipeline that tweaks the geometry of a phantom model to match the pathology seen in real radiographs, such as those afflicted by RA. This will allow clinicians access to clean (already segmented) volumetric rendering of pathological skeletal anatomy without the high doses of ionising radiation typically associated with CT scans, or where there is no access to volumetric scans.

In the next chapter, we are planning to address some of the problems discussed in previous section:

- We can pad white spaces (same pixel values as background) around images to make them bigger. The alternative way is to impose constraints on parameters where fingers are not allowed to extend outside images.

- There are some images (e.g. Image 5) where fingers are mis-matched or overlapped. We can look into multi-objective optimisation algorithms with both MAE and ZNCC as objectives. ZNCC is very helpful for shape matching. The algorithm are working by minimising MAE and maximising ZNCC at the same time.

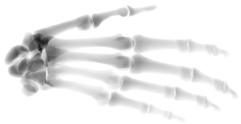
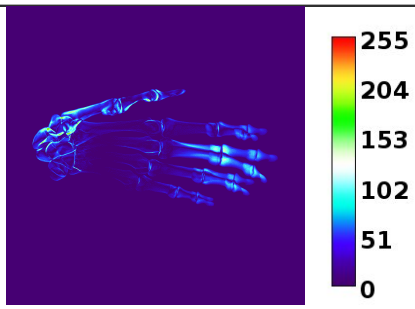
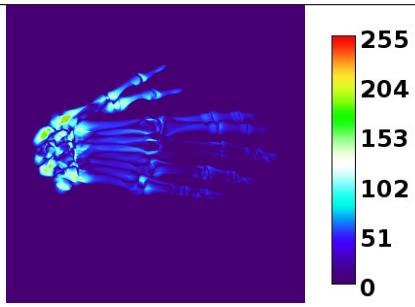
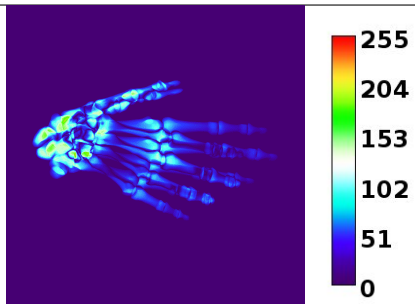
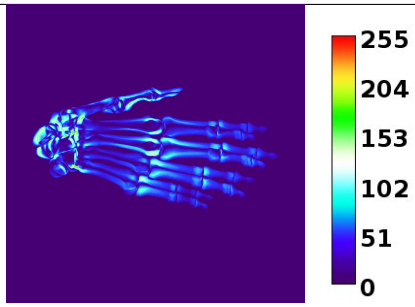
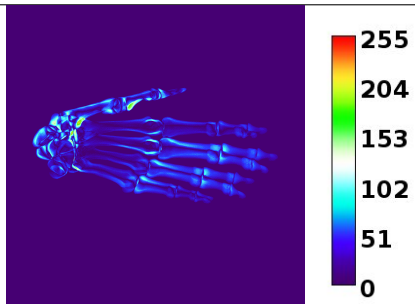
Methods	X-ray image	Error map
Nelder-mead MAE = 3.45E-4		
CG MAE = 5.41E-4		
PRS MAE = 5.62E-4		
SA MAE = 3.85E-4		
EA MAE = 3.17E-4		

Figure 3.6: *Comparing individual result for each method on the hand registration problem (median solution over 15 runs, i.e. solution corresponding to the green bars in Figure 3.5).*

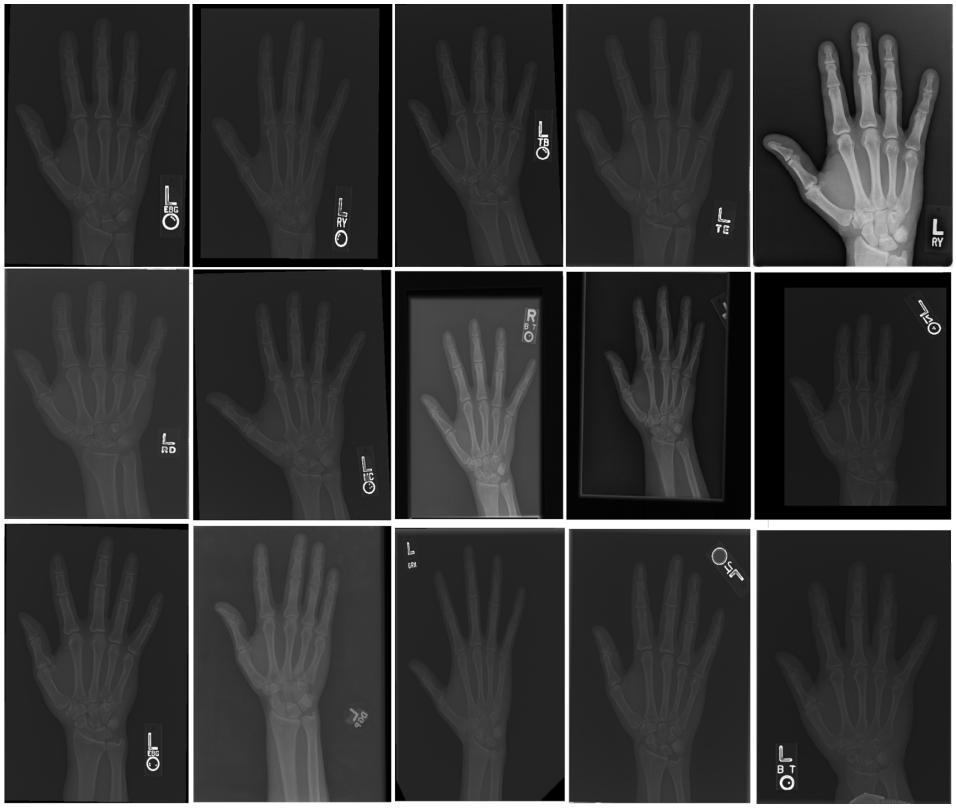


Figure 3.7: *15 selected hand radiographs taken from the MURA dataset without data processing. Images are numbered from image 1 (top left) to image 15 (bottom right). Top row: image 1-5, middle row: image 6-10, and bottom row: image 11-15.*

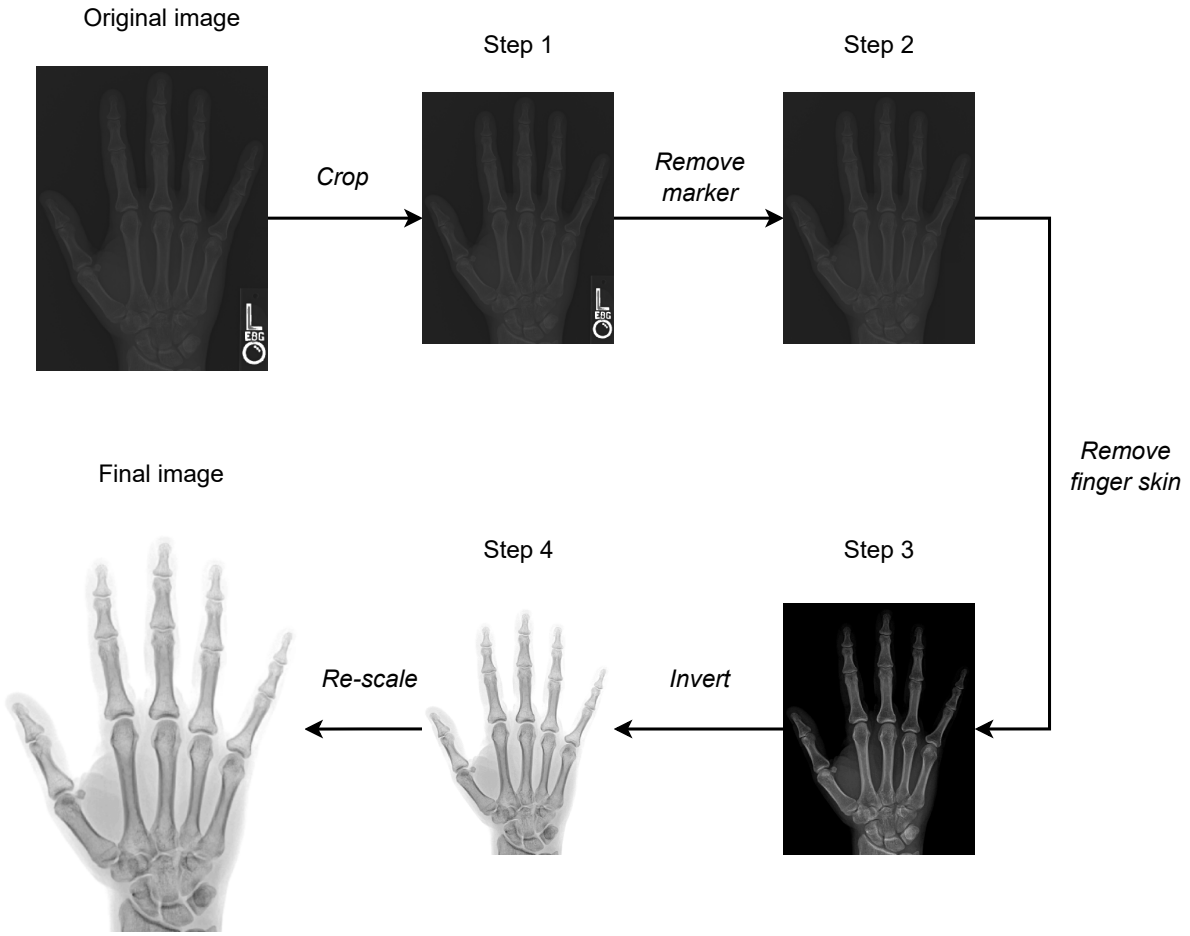


Figure 3.8: *Illustrations of image pre-processing.*

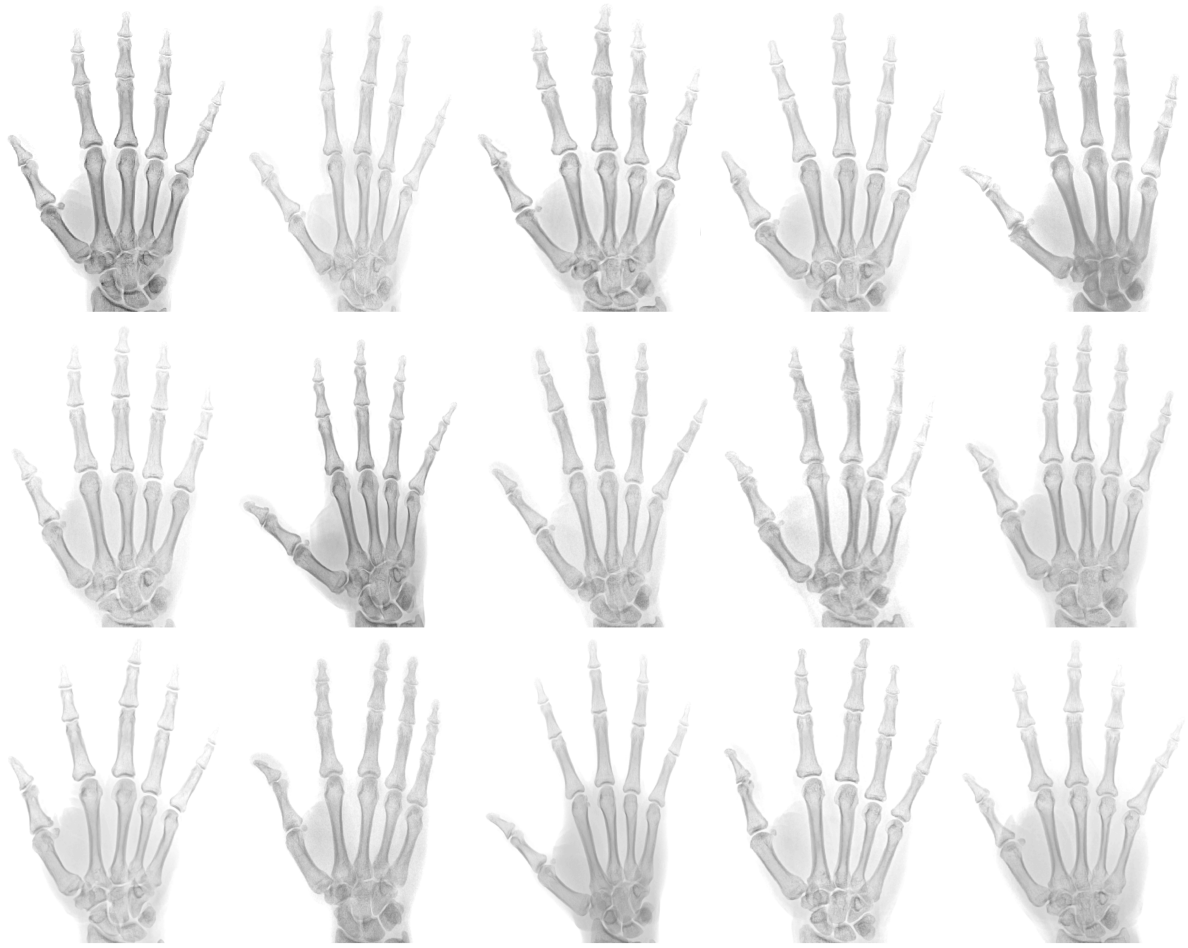
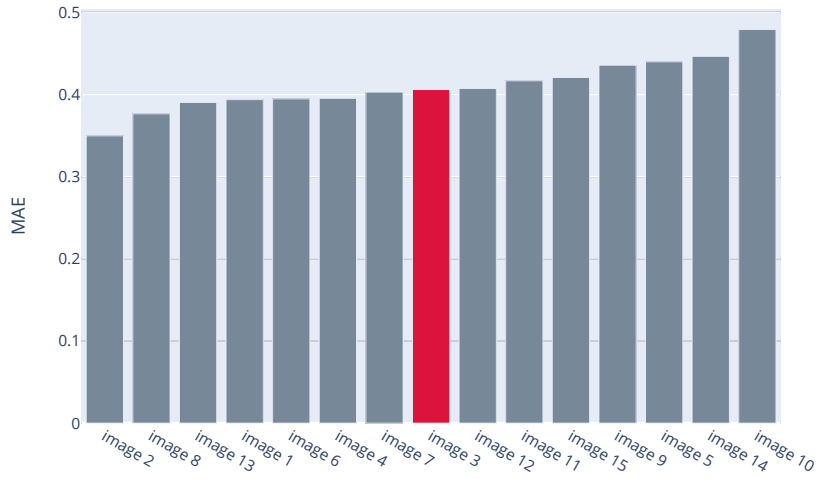
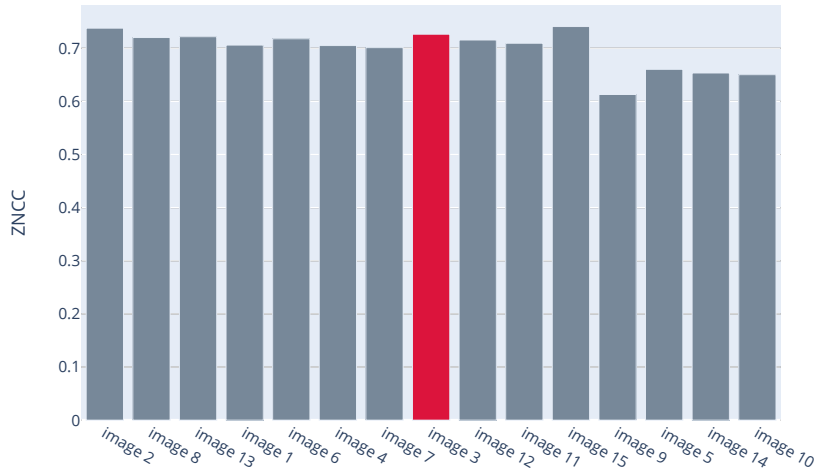


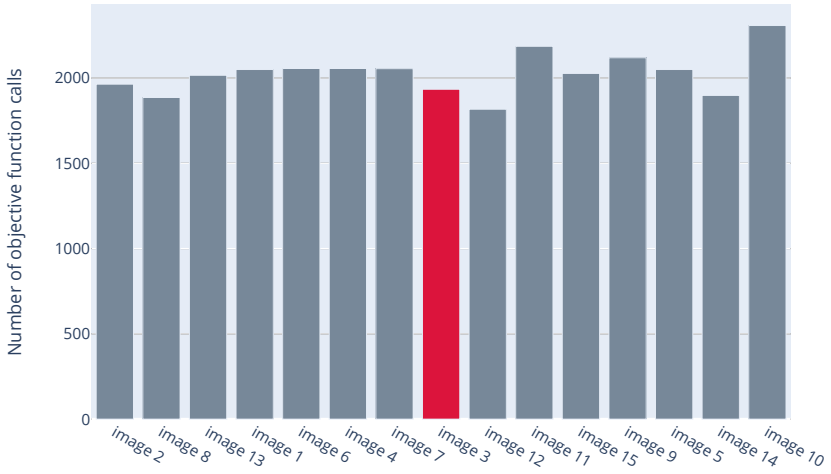
Figure 3.9: *15 selected hand radiographs from the MURA dataset after pre-processing. Images are numbered from Image 1 (top left) to Image 15 (bottom right). Top row: Images 1-5, middle row: Images 6-10, and bottom row: Images 11-15.*



(a) *MAE of each registration*



(b) *ZNCC of each registration*



(c) *Number of objective function calls of each registration*

Figure 3.10: *Bar charts for quantitative results shown in Table 3.3. All data is sorted on MAE and the median result is highlighted in red.*

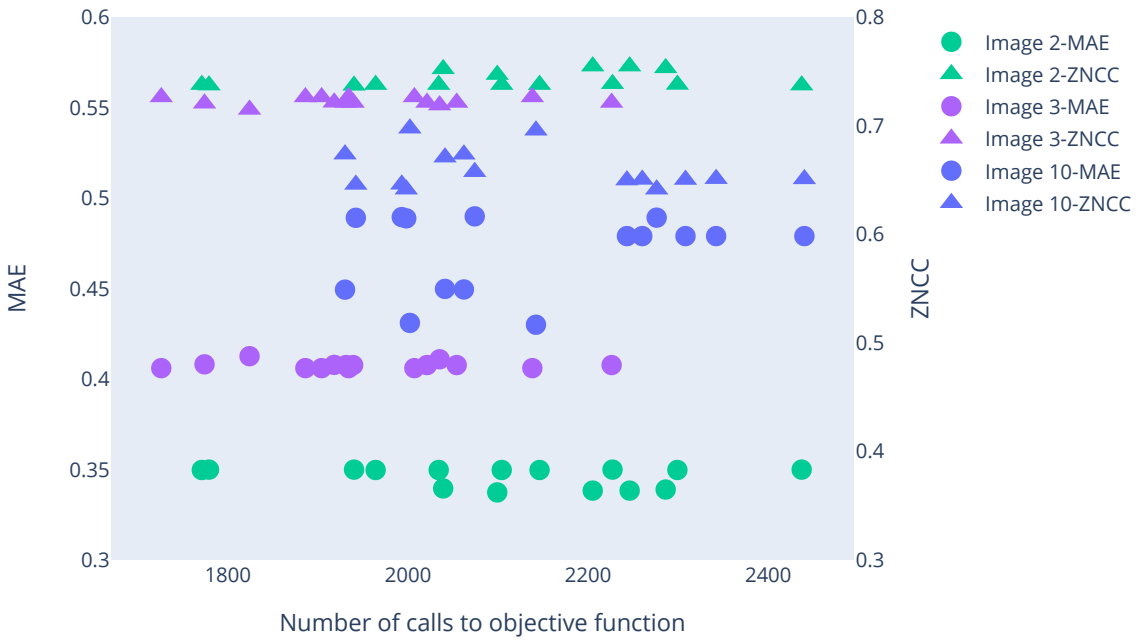


Figure 3.11: *Scatter plots of results for 15 registrations on Image 2, 3 and 10. Circles represent MAE, triangles represent ZNCC and different colours represent different images.*

Chapter 4

Registration of real radiographs using bi-objective optimisation algorithms

4.1 Introduction

The previous chapter demonstrated that the registration of a 3D hand model to real clinical X-ray images can be performed. We started with a simple test case where synthetic X-ray images are used. This is helpful in developing our method where different optimisation algorithms are investigated. Among those optimisation algorithms, EA performed the best in terms of accuracy and stability, which is a good choice to include in our registration method. Then we progressed to registration of real clinical radiographs using CMA-ES, a type of EA that aims to solve difficult non-linear black-box optimisation problems. Registration performance is good in most cases where fingers are recovered.

Unlike many other registration methods [4], [58], [127], [128], [154], [175], our method only requires a single X-ray image. It is very helpful in medical applications. For example, diagnosis of hand disease such as RA often requires multiple X-ray images of hands that are taken by radiologists. The patient can be exposed to high level of X-rays. Our method brings the analysis of a single 2D projection into 3D space. Our approach rely on fast X-ray generation (by gVirtualXRay) and optimisation algorithms to tune deformation parameters that transform the 3D hand model to finely match input images.

In this chapter we investigate the effect of integrating a bi-objective optimisation algorithm into our registration framework. A bi-objective optimisation is a special case where two metrics are used to construct the objective function. Multi-objective optimisation has been applied in solving many engineering problems such as flight planning [43], building design [176] and agricultural system management [174]. Recently, multi-objective optimisation is also applied in image registration problems [28], [114], [135]. In a multi-objective optimisation problem, multiple metrics are being optimised at the same time. As we discussed in previous chapters one metric is often not enough for complex registration problems such as our hand registration problem. However, as more metrics are computed at each iteration during the optimisation process, more computational resources are required. We use NSGA-II along with two conflicting metrics to solve our hand registration problem.

The main contribution of this chapter is the use of bi-objective optimisation algorithms to address deficiencies of our proposed method, i.e. we aim to register images that were not registered successfully in the previous chapter. Also, we compare the registration performance between single- and bi-objective optimisation algorithms.

The rest of the chapter is organised as follow: Section 4.2 describes our methods and techniques about the bi-objective optimisation. Section 4.3 describes our results as well as quantitative and qualitative evaluation. Section 4.4 provides conclusions and describes future work.

4.2 Bi-objective optimisation

We use NSGA-II as the optimisation algorithm. NSGA-II is a multi-objective EA [32]. Further details about NSGA-II can be found in Section 2.5.6. Before the optimisation, all images are normalised to have zero-mean and unit-variance in the same way as in previous experiments. We use MAE and ZNCC to construct the objective function, i.e. as the fitness function to be optimised by NSGA-II. MAE is the sum of absolute errors between samples and then divided by total number of samples (Eq. 2.19). The best number can be achieved is zero. Typically, lower MAE value indicates better optimising result. ZNCC is a measurement of similarity

between two images. To compute ZNCC (see Eq. 2.2), the target and predicted images are normalised first, which is subtracting all pixels by the mean value and divided by standard deviation. Normalised target and predicted images are then multiplied. Finally, all values are added and divided by total number of pixels. ZNCC primarily concentrates on template matching. Further details about ZNCC can be found in Section 2.4.1.

Our registration problem is considered to be complex both in terms of number of parameters and the corresponding data range. Full details about parameters and their ranges were presented in Table 3.2

Typically, in a single-objective optimisation problem, the solution with best objective value is chosen. For example, it can be a solution with the lowest MAE or a solution with the highest ZNCC. However, it is not possible in the bi-objective optimisation. There is always a scenario where one solution have the lowest MAE) but their ZNCC is lower than others. To compare the performance of CMA-ES and NSGA-II, we need to determine which solution is most suitable to reflect the performance of NSGA-II. There are some performance measurements that measure the quality of a PF [38], [107], [152]. However, it is not useful in our case since we use different images in the bi-objective optimisation. Generally speaking, solutions with with extreme objective values are not suitable since they tend to provide extremely good or bad results. The solution with median objective value is more suitable since they have more balance between different objectives.

4.3 Results

Two metrics are used in the objective function: MAE and ZNCC. MAE is a dissimilarity metric and ZNCC is a similarity metric. Since one can minimising MAE or maximising ZNCC, there should be a way to optimising both. In our case we minimise both MAE and the negative of ZNCC. NSGA-II is used as the optimisation algorithm. The experiment set-up is the same as discussed in that experiment except there is no restart mechanism in this experiment. Intuitively, NSGA-II are expected to perform not worse than CMA-ES and we observed that the objective

is not improving after 500 iterations from CMA-ES. We set number of iterations at 1,000 in NSGA-II. If NSGA-II can have similar performance using only half of the computing time than CMA-ES and have no restart, then NSGA-II is preferred to be included in our registration framework.

4.3.1 Selection of individual in a Pareto-front

To compare with results obtained from the previous experiment the following solutions are selected:

- The solution with the lowest MAE (NSGA-II).
- The solution with the highest ZNCC (NSGA-II).
- A solution with median MAE and ZNCC (NSGA-II).
- The solution with the lowest MAE (CMA-ES).
- The solution with the highest ZNCC (CMA-ES).

In Figure 4.1 we plot the PF at the last iteration from NSGA-II and 5 solutions listed above. For better visualisation and comparison, 5 images (3 from NSGA-II and 2 from CMA-ES) for each registration (15 in total) are displayed. In all 15 registrations, three chosen individuals provided similar results (Figure 4.1). Selecting the solution with median objectives is appropriate in our case. Next, we use the solution with median objectives in NSGA-II to compare with the best solution in CMA-ES.

4.3.2 Computational performance

Figure 4.2 shows two objectives against number of iterations for both CMA-ES and NSGA-II for 15 registrations. Results from CMA-ES are restated and only MAE is used to construct the objective function. ZNCC is computed between final predicted images and reference images.

Both MAE and ZNCC are used to construct the objective function in NSGA-II. The average value of MAE at each iteration is computed and is represented by red solid line in Figure 4.2. The average value of ZNCC at each iteration is computed and is represented by red dotted line in Figure 4.2. As shown in the figure both red and black lines are flat, i.e. both algorithms has converged. Thus, it is not unfair to compare two algorithms directly since sufficient number of iterations (running time) are given.

NSGA-II has achieved higher ZNCC (red dotted line in Figure 4.2) in almost all registrations (except 14) than CMA-ES. NSGA-II has achieved lower MAE (red line in Figure 4.2) in three registrations (4,10,11). In Registration 10 NSGA-II has clearly outperformed CMA-ES with marginal improvements of MAE and ZNCC. Also, by looking at the change of slope of lines (black and red), NSGA-II tends to converge faster and requires lower number of iterations. Overall, NSGA-II has better performance than CMA-ES with the benefit of fast convergence and better (or similar) objective values.

4.3.3 Visual performance

In Figure 4.1, points located at the top left corner is considered better since they have lower MAE and higher ZNCC than points located at the right bottom corner since they have higher MAE and lower ZNCC. Based on this observation, there are three registrations (4,10,11) that NSGA-II should perform better than CMA-ES. All points from NSGA-II are located at the top corner and two points from CMA-ES are located at the right bottom corner. By comparing actual predicted images, NSGA-II fixed the problem of finger overlapping in Registration 4 and fixed the problem of finger mis-matching in Registration 10. In Registration 11, two algorithms have similar performance. Generally speaking, individuals with lower MAE and higher ZNCC tends to have better results. However, it is not always the case. For example, CMA-ES should perform better than NSGA-II in Registration 14. The results are not visible in the figure. Better images are provided in Table 4.1. From the table NSGA-II is actually performing better than CMA-ES where the problem of finger extended outside of the image is fixed.

Table 4.1 displays 6 images per row: the target image, the predicted image with lowest MAE from CMA-ES, the predicted image with highest ZNCC from CMA-ES, the predicted image with lowest MAE in a PF from NSGA-II, the predicted image with highest ZNCC in a PF from NSGA-II and the predicted image with median MAE and ZNCC in a PF from NSGA-II. From the table, we can clearly see that there are 4 registrations (4,10,12,14) that NSGA-II produces better results comparing to CMA-ES. NSGA-II have similar performance comparing to CMA-ES in the rest of registrations. Overall, NSGA-II has better performance than CMA-ES in terms of qualities of predicted images. We also notice that both algorithms seem to have wrong alignments in some registrations, e.g. in Registration 5, the fingers are all moving towards to left hand side. One possible reason is that the hand might not be centred in the target image. For example, CMA-ES provided similar result in Registration 5 and 10 while NSGA-II corrected the wrong alignment in Registration 10 but not in Registration 5 (row “Image 5" and “Image 10" in Table 4.1). The hand is appearing to be more centred in Image 10 comparing with Image 5. Further investigation is necessary to identify the cause of this problem.

Table 4.1: *Display target images and predicted images from CMA-ES and NSGA-II. Images are the same as displayed in Fig. 4.1 (excluding target images).*

Target	**CMA-ES best MAE**	**CMA-ES best ZNCC**	**NSGA-II best MAE**	**NSGA-II best ZNCC**	**NSGA-II Pareto Front**

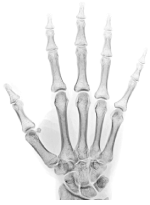
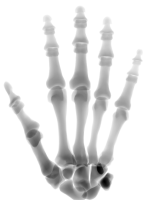
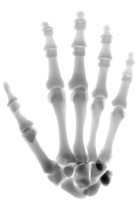
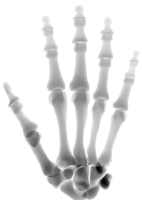
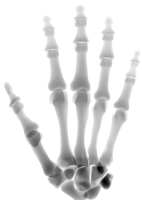

				

Image 1					
MAE=	0.3937	0.4099	0.4005	0.4045	0.4023
ZNCC=	0.7060	0.7137	0.7173	0.7216	0.7207

Continued on next page

Table 4.1 – continued from previous page

Target	CMA-ES best MAE	CMA-ES best ZNCC	NSGA-II best MAE	NSGA-II best ZNCC	NSGA-II Pareto Front

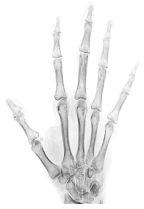
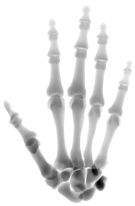
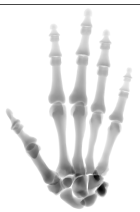
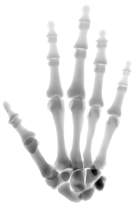
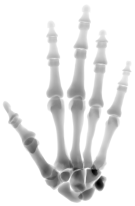
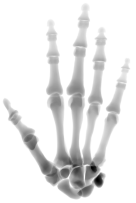

					

Image 2					
MAE=	0.3497	0.3540	0.3503	0.3523	0.3511
ZNCC=	0.7378	0.7405	0.7457	0.7474	0.7471

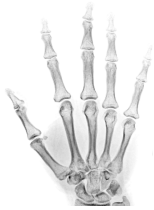
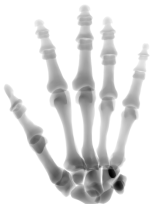
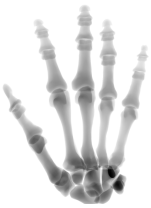
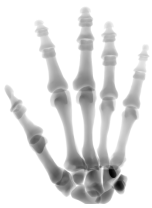
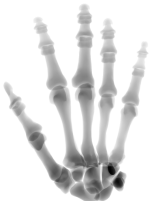
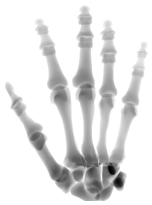

					

Image 3					
MAE=	0.4060	0.4068	0.4071	0.4094	0.4083
ZNCC=	0.7264	0.7269	0.7283	0.7298	0.7289

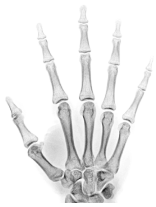
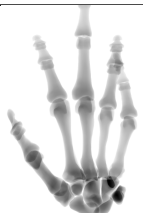
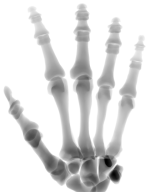
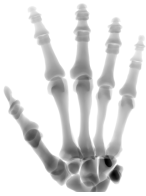
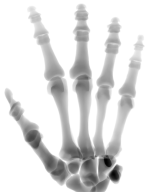

					

Image 4					
MAE=	0.3953	0.3991	0.3912	0.3940	0.3922
ZNCC=	0.7051	0.7056	0.7115	0.7126	0.7123

Continued on next page

Table 4.1 – continued from previous page

Target	CMA-ES best MAE	CMA-ES best ZNCC	NSGA-II best MAE	NSGA-II best ZNCC	NSGA-II Pareto Front

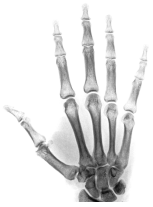
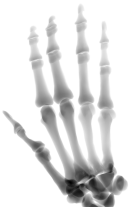

		

Image 5					
MAE=	0.4398	0.4449	0.4401	0.4437	0.4417
ZNCC=	0.6603	0.6639	0.6613	0.6667	0.6657

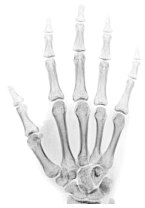
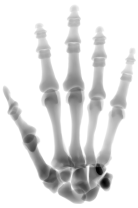
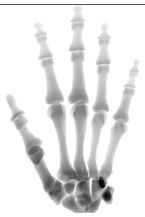

			

Image 6					
MAE=	0.3947	0.3952	0.3968	0.4022	0.4010
ZNCC=	0.7179	0.7183	0.7209	0.7348	0.7342

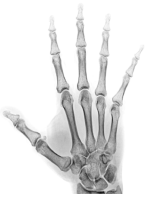
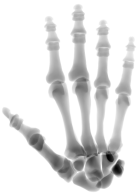
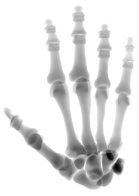
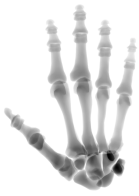
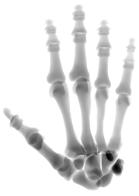
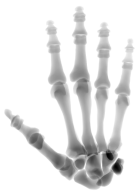

					

Image 7					
MAE=	0.4029	0.4064	0.4037	0.4080	0.4062
ZNCC=	0.7012	0.7079	0.7080	0.7146	0.7131

Continued on next page

Table 4.1 – continued from previous page

Target	CMA-ES best MAE	CMA-ES best ZNCC	NSGA-II best MAE	NSGA-II best ZNCC	NSGA-II Pareto Front

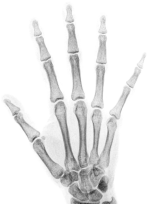
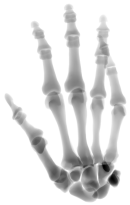
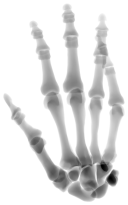
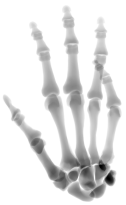

					

Image 8					
MAE=	0.3765	0.3780	0.3791	0.3822	0.3809
ZNCC=	0.7202	0.7223	0.7315	0.7338	0.7318

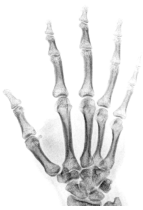



Image 9					
MAE=	0.4356	0.4446	0.4387	0.4440	0.4410
ZNCC=	0.6130	0.6229	0.6272	0.6349	0.6332

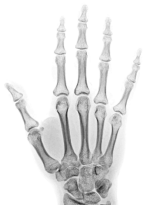
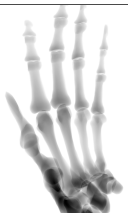
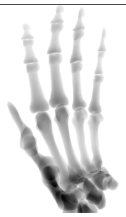
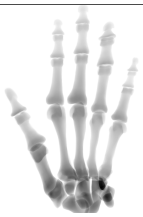
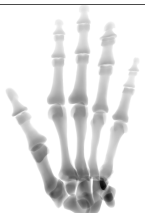
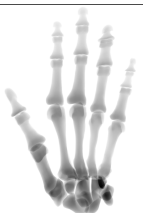
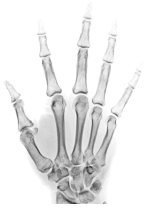
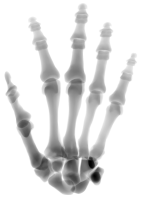
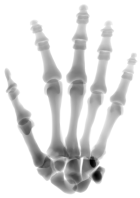
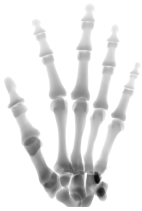
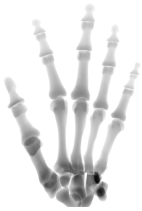
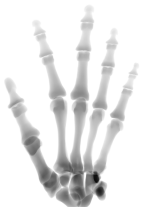
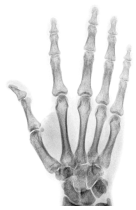
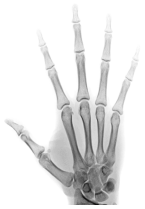
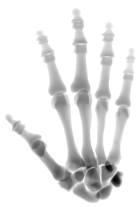
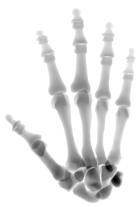

					

Image 10					
MAE=	0.4790	0.4935	0.4423	0.4480	0.4459
ZNCC=	0.6503	0.6529	0.6909	0.6958	0.6950

Continued on next page

Table 4.1 – continued from previous page

Target	CMA-ES best MAE	CMA-ES best ZNCC	NSGA-II best MAE	NSGA-II best ZNCC	NSGA-II Pareto Front
 Image 11					
MAE=	0.4167	0.4211	0.4108	0.4136	0.4118
ZNCC=	0.7094	0.7146	0.7315	0.7346	0.7337
 Image 12					
MAE=	0.4073	0.4083	0.4115	0.4137	0.4123
ZNCC=	0.7156	0.7171	0.7237	0.7256	0.7250
 Image 13					
MAE=	0.3903	0.3910	0.3911	0.3932	0.3923
ZNCC=	0.7218	0.7237	0.7261	0.7283	0.7275

Continued on next page

Table 4.1 – continued from previous page

Target	CMA-ES best MAE	CMA-ES best ZNCC	NSGA-II best MAE	NSGA-II best ZNCC	NSGA-II Pareto Front

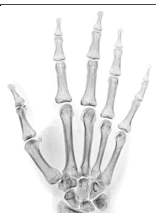
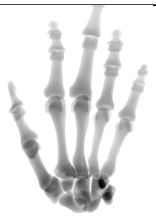
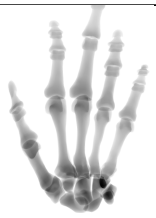
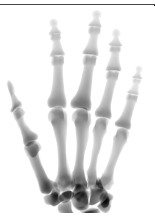
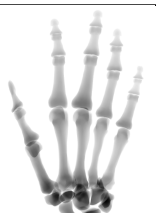
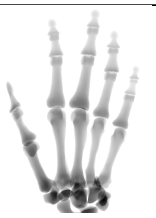

					

Image 14					
MAE=	0.4464	0.4478	0.4716	0.4757	0.4740
ZNCC=	0.6535	0.6547	0.6449	0.6513	0.6502

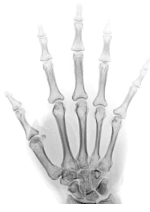



Image 15					
MAE=	0.4206	0.4322	0.4223	0.4295	0.4259
ZNCC=	0.7409	0.7452	0.7424	0.7527	0.7458

4.4 Discussion and conclusion

We investigated the effect of bi-objective optimisation that aimed to improve the registration accuracy. Registrations of 15 different radiographs were performed in the bi-objective optimisation where NSGA-II was used. The objective function was constructed by using two conflicting metrics: MAE and ZNCC. We investigated how to select the best solution from PF in the context of the bi-objective optimisation. In our hand registration problem, the choice of the best solution have little impact on final results, i.e. all predicted images from those solutions are similar to each other. We then compared the performance of single- and bi-objective optimisation algorithms using both computational (objective values and convergence of optimisation) and visual (the quality of predicted images) comparisons. Use of

NSGA-II greatly improves the registration performance where lower MAE and higher ZNCC were recorded and faster convergence of the optimisation in most cases. NSGA-II has achieved better or similar results on all test images comparing to CMA-ES from the visual comparison. Overall, NSGA-II is a better choice over CMA-ES in solving our hand registration problem. One particular problem observed is that there are similar wrong alignments of fingers in both registrations (e.g. in Registration 5) using different optimisation algorithms. It indicates that the problem might lies on the target images. Further investigation is necessary to identify the cause of the problem.

In the next chapter we plan to investigate why some images are harder to register than other by conducting manual registrations, i.e. manually tuning the transformation parameters so that the 2D projection is finely matched to input image. If two images can be matched manually, our expectation is that optimisation algorithms are able produce good results as well. We also plan to extend our method using another dataset that involves different musculoskeletal anatomy to test the adaptability of our method.

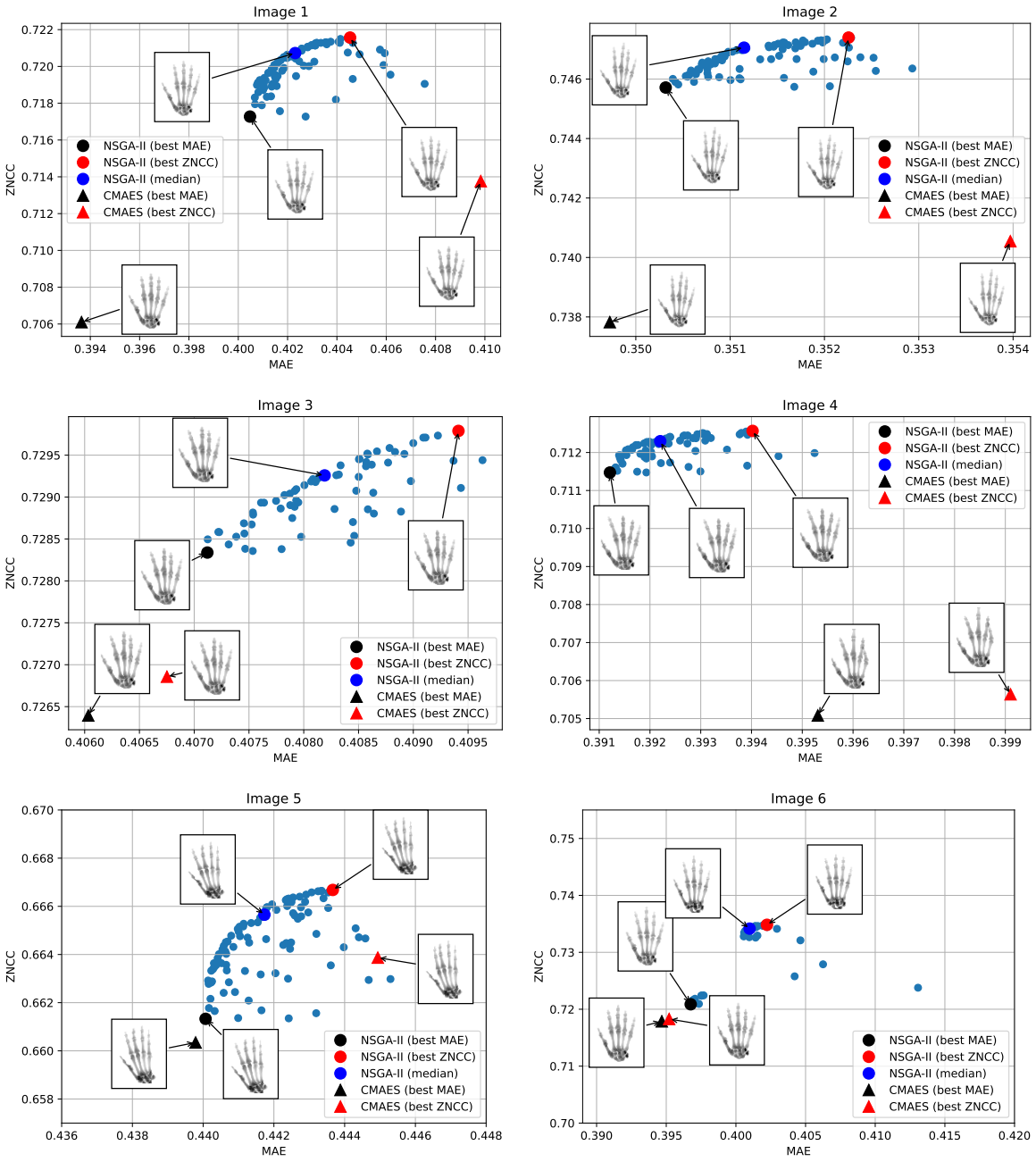


Image 1
Image 2
Image 3
Image 4
Image 5
Image 6
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)
ZNCC
MAE

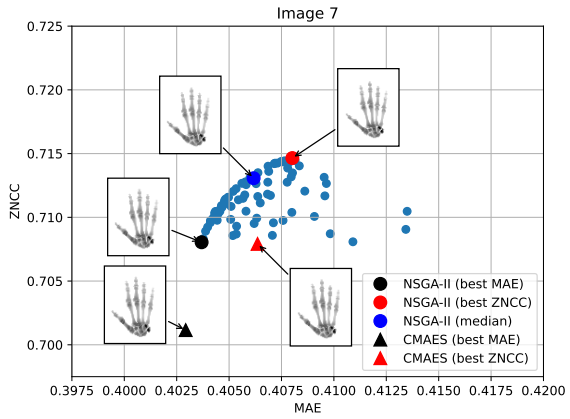


Image 7
ZNCC
MAE
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)

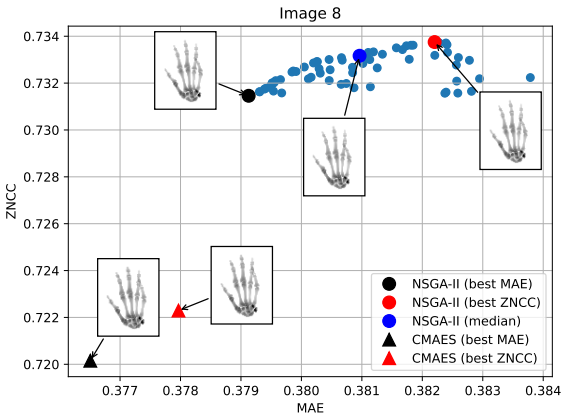


Image 8
ZNCC
MAE
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)

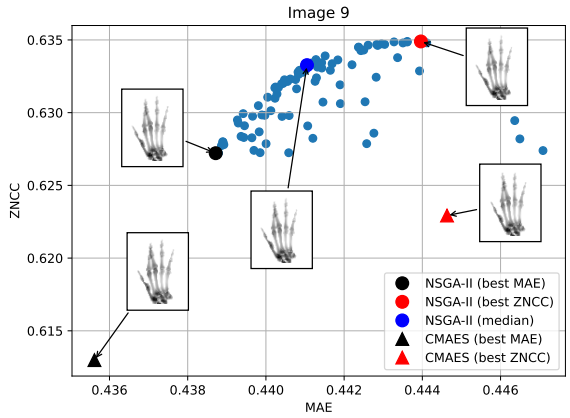


Image 9
ZNCC
MAE
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)

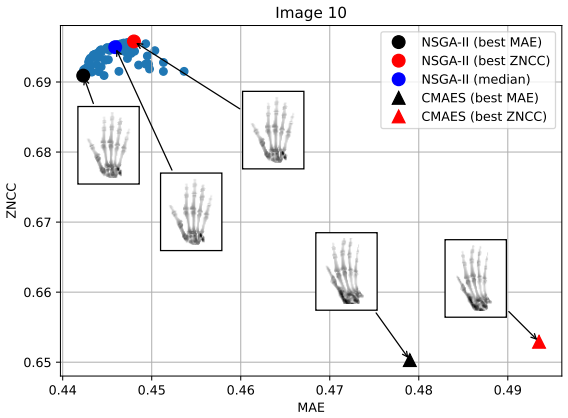


Image 10
ZNCC
MAE
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)

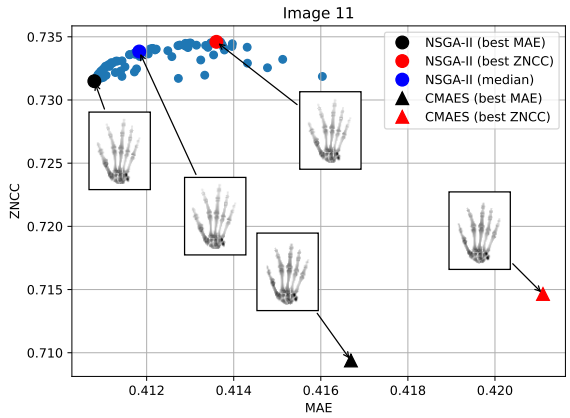


Image 11
ZNCC
MAE
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)

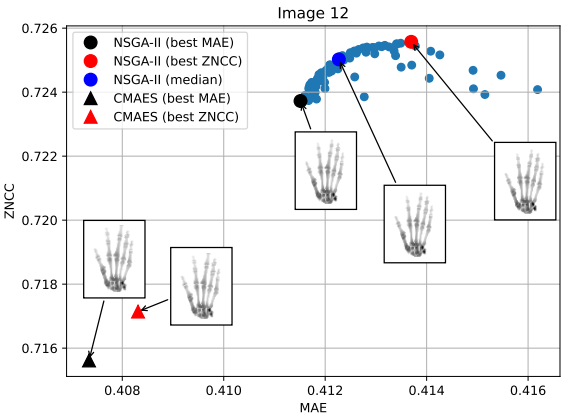


Image 12
ZNCC
MAE
NSGA-II (best MAE)
NSGA-II (best ZNCC)
NSGA-II (median)
CMAES (best MAE)
CMAES (best ZNCC)

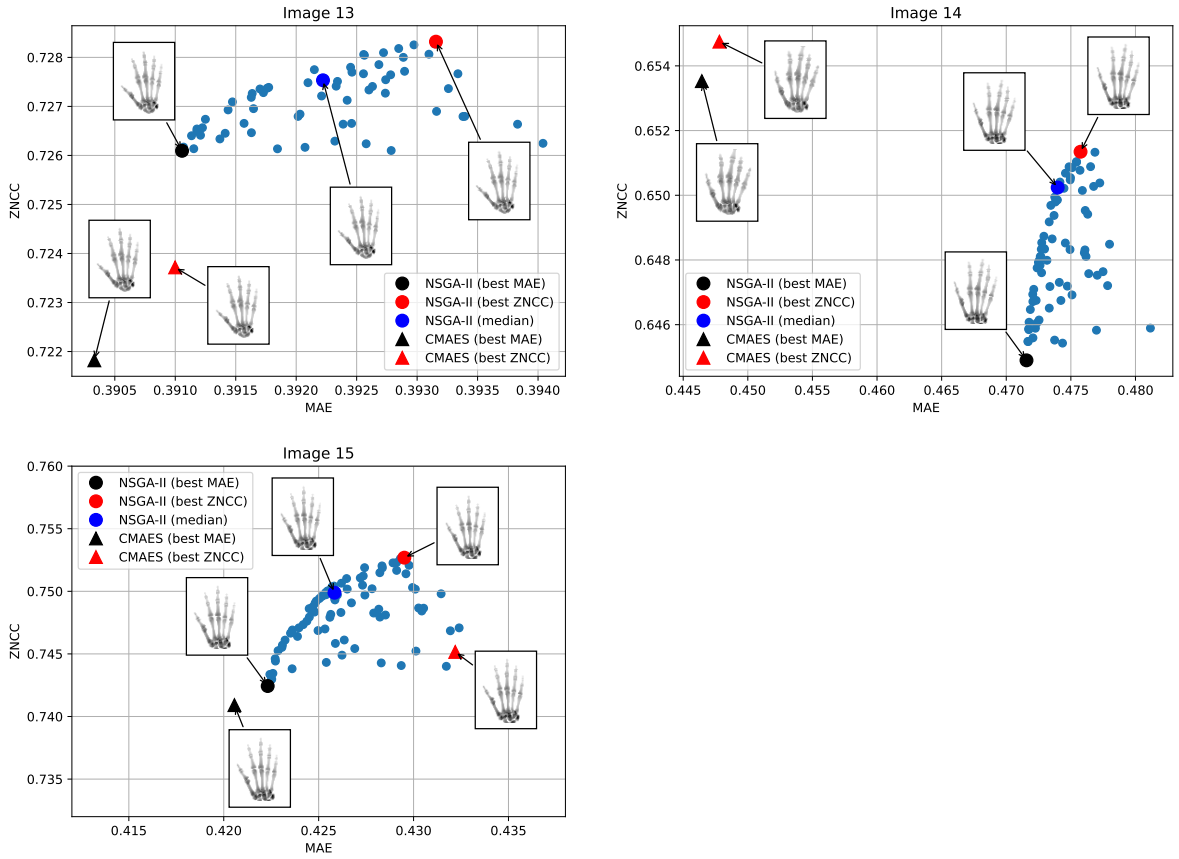


Figure 4.1: *Plots of results from CMA-ES and NSGA-II for 15 images. Image displayed: two individuals with best metric value (lowest MAE and highest ZNCC) and one individual with median metric value in a PF from NSGA-II. Two individuals with best metric value (lowest MAE and highest ZNCC) from CMA-ES. Individuals with lowest MAE and highest ZNCC are located in the top left corner in the figure.*

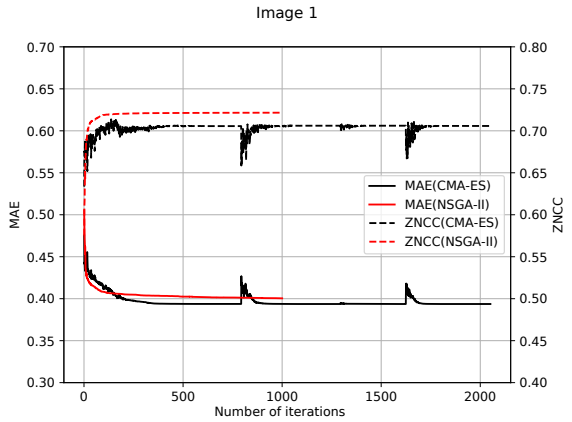


Image 1
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

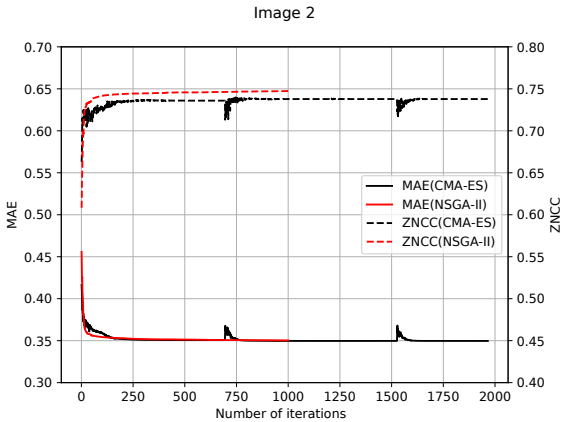


Image 2
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

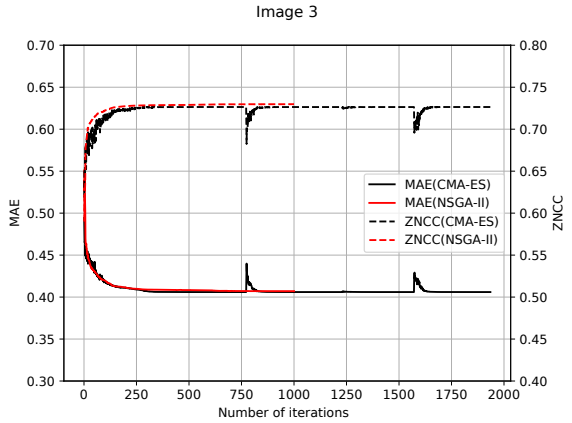


Image 3
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

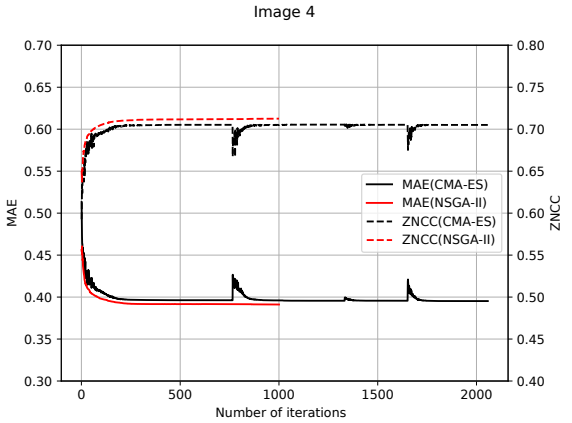


Image 4
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

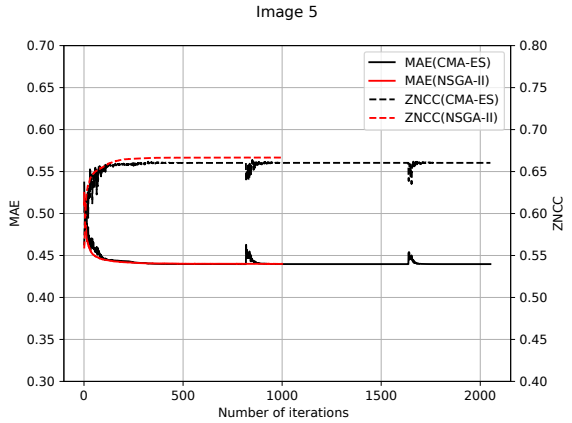


Image 5
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

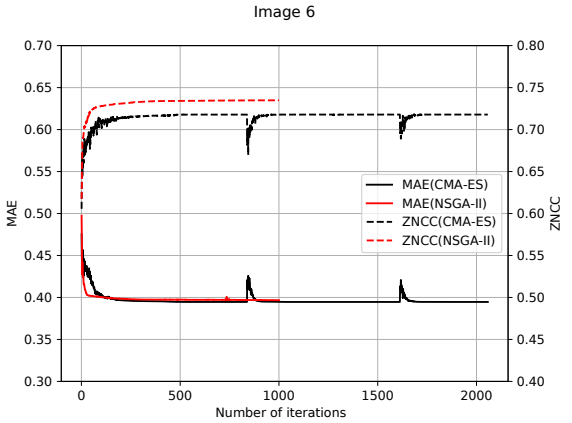


Image 6
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

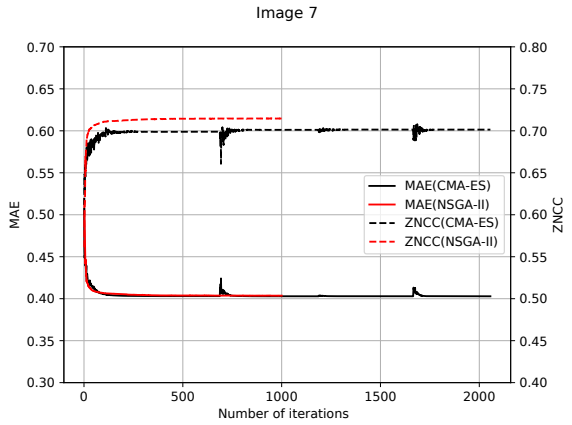


Image 7
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

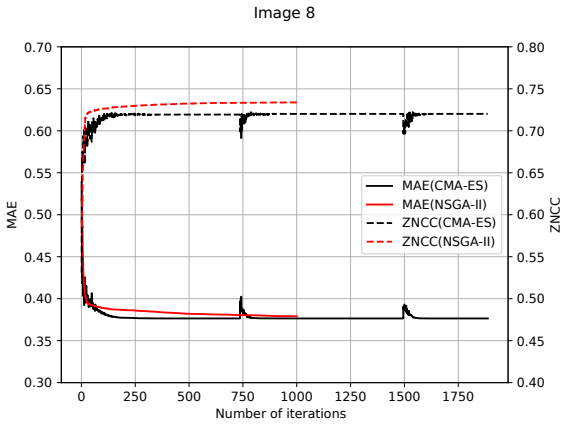


Image 8
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

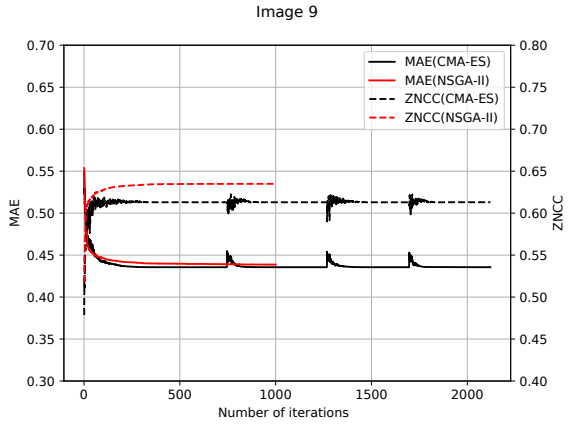


Image 9
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

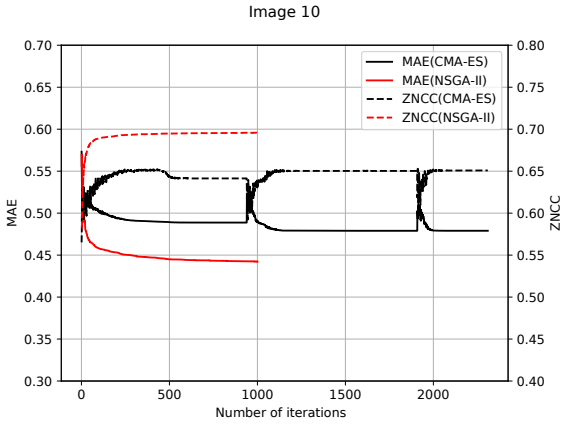


Image 10
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

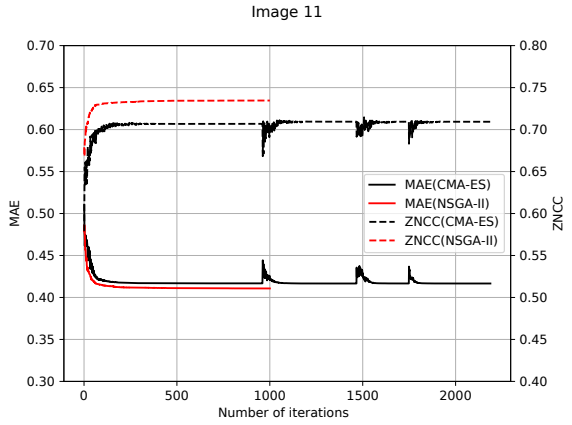


Image 11
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

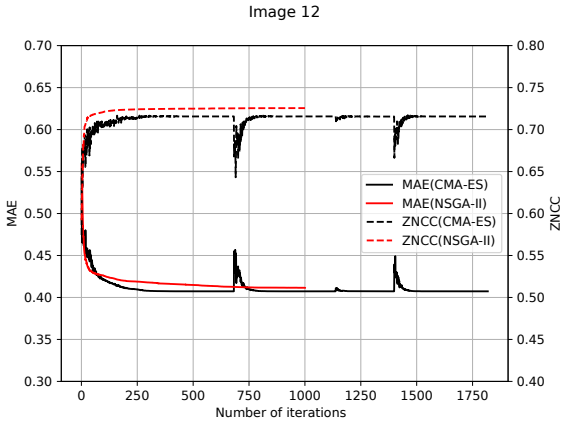


Image 12
MAE
ZNCC
Number of iterations
MAE(CMA-ES)
MAE(NSGA-II)
ZNCC(CMA-ES)
ZNCC(NSGA-II)

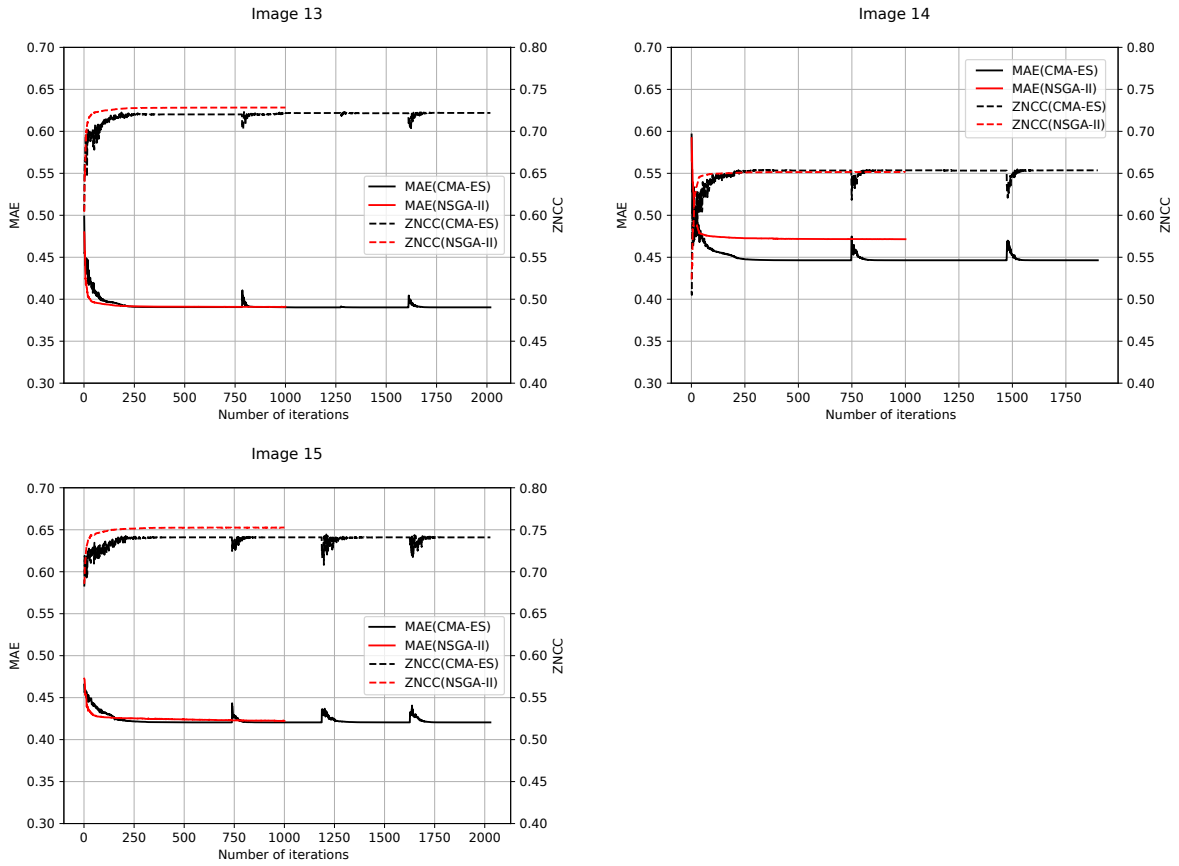


Figure 4.2: *Plot MAE and ZNCC at each iteration for both CMA-ES and NSGA-II. The dotted line represents the change of ZNCC over iterations. The solid line represents the change of MAE over iterations. CMA-ES is marked using black colour and NSGA-II is marked using red colour. In CMA-ES a single objective function is called at each iteration. In NSGA-II two objective functions are called at each iteration and the average value of two metrics in a Pareto-front is computed at each iteration.*

Chapter 5

Registration of real radiographs using multi-objective optimisation algorithms

5.1 Introduction

In the previous chapter we investigated the effect of two techniques including image padding and bi-objective optimisation aimed to improve the registration accuracy using fifteen test-cases. The use of image padding could affect the registration accuracy to some degree. However, there was no clear indication that the registration accuracy is improved after padding. Also, more computational resources were required to complete the registration where it might not be feasible in some cases. The use of NSGA-II generally had positive impact on registration results, i.e. NSGA-II tends to produce better or similar results compared with CMA-ES on all fifteen test-cases. However, three radiographs were not registered successfully, which demonstrated the complexity of our registration problem.

This chapter concentrates on further improvements of our registration framework. Firstly, we conduct manual registrations to investigate why some images are harder to register than others. Manual registration is the process of matching target and predicted images by tuning transformation parameters by hand. In this way, based on information that is obtained from the manual registrations, different techniques can be applied to improve our registration method. As discussed in the previous chapter, both CMA-ES and NSGA-II tend to produce similar wrong alignment in registering same images. This might indicates that the optimisation algorithms is

not the problem for failed registration. The problem might lies on the input images, i.e. hands in both images might not be able to match due to their position in the image. Through manual registrations the underlying cause can be identified.

Also, we adapt our registration framework to another registration problem, which is the registration of a 3D model to the hip radiograph with Anterior Posterior (AP) views. This helps us to further test the flexibility and general adaptability of our registration framework. In previous chapters we mainly focus on advances of optimisation algorithms which improved registration accuracy in some cases. In this chapter different similarity and dissimilarity metrics would be evaluated in the context of multi-objective optimisation, which can help us to identify suitable metrics that can be used together in solving the hand registration problem.

The main contribution of this chapter is enabling the use of manual registration to understand the causes of problems encountered in previous experiments (Chapters 3 and 4). The findings are used to address deficiencies of our registration method and for further development of our method. We also investigate the use of multi-objective optimisation algorithms and the relationship between different metrics.

The rest of the chapter is organised as follow: Section 5.2 discusses the manual registration process, another dataset and the multi-objective optimisation, Section 5.3 presents registration results for both hand and hip registrations and Section 5.4 summaries our findings.

5.2 Methods and materials

5.2.1 Manual registration

Manual registration is a way to investigate why some images are harder to register than others. In the manual registration process, parameters are tuned by a human operator where the 3D hand model is transformed to match the input image. Each manual registration takes around 5 minutes. We concentrate on three images (5, 8 and 9 as shown in Figure 3.9) where registrations

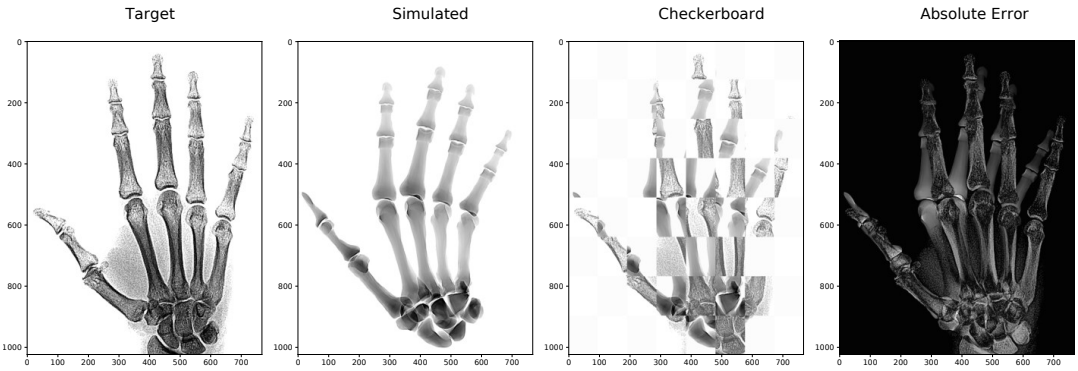


Figure 5.1: *The alignment of hands (Image 5) by tuning transform parameters manually instead of using optimisation algorithms.*

of hands were not successful. An example of manual registration of Image 5 is shown in Figure 5.1. All fingers except the thumb are not able to be registered. We also conduct manual registrations for both Images 8 and 9. Similar problems are observed and our main findings are summarised as follow:

1. The rotation of the whole hand is not able to recover the position of the wrist. This is because the wrist of the hand is not centred in the target image which results in different position of hands. Thus, the registration of other part of the hand is not possible.

2. The boundary ranges of some parameters are not wide enough. We determined the boundary range based on real-life observations of finger movements (see Table 3.2). However, during manual registration we identified that some fingers actually need more rotation to recover the correct position. This might due to the fact that clinical hand dataset we used might contains abnormal radiographs.

3. The edge of finger bones is not as sharp as other bones. This might be caused by removal of skins during pre-processing of raw images. By highlighting the edges of bones, the chance of successful registration is greater.

To solve these problems, we consider three techniques:

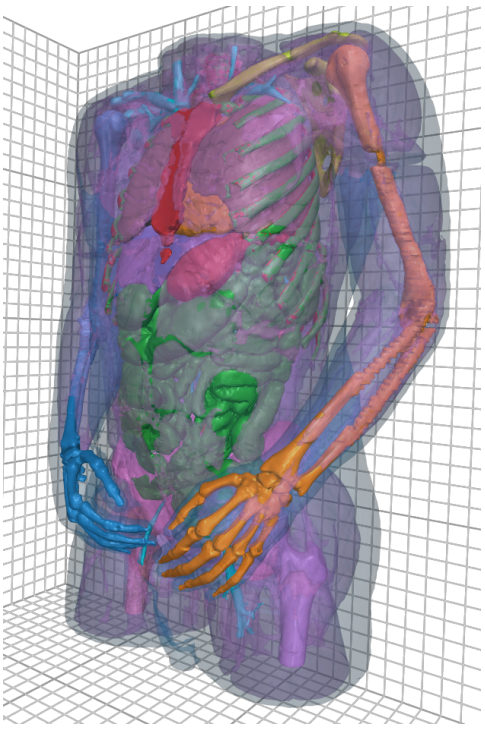


Figure 5.2: *Surface models extracted from the labelled data (taken from cadaver).*

1. We can make use of image padding technique. Instead of padding around the image, we only pad images on the right hand side so that the wrist is roughly centred in the target image.

2. We can ease the boundary range for particular parameters.

3. We can add a sharpening filter where details are enhanced. A sharpening filter is a technique to highlight area of interest and enhance those regions.

5.2.2 Visible Human Project (VHP)

The *Visible Human Project (VHP)* is created by National Library of Medicine (NLM) that aimed to provide "publicly-available complete, anatomically detailed, three-dimensional representations of a human male body and a human female body" [1]. The male dataset, the *Visible Man* was released in 1994. The female dataset, the *Visible Woman* was released in 1995. The Visible Human Male data set includes MRI, CT, X-ray radiographs (Figure 5.3) and

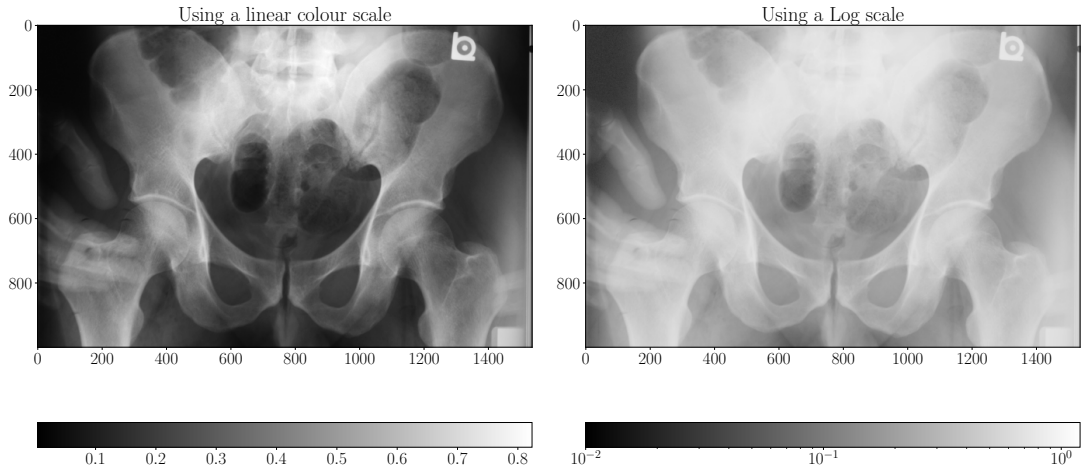


Figure 5.3: *Radiographs with AP pelvis view (taken alive) from the Visible Man. The linear scale is used on the left hand side and the log scale is used on the right hand side.*

cross-sectional cryosection, which were made from the male's cadaver. *Voxel-Man's* Segmented Inner Organs (SIO) is a model composed of a set of segmented volumetric images obtained from the original cross-sectional cryosection [56], [115]. The 3D surface model can be created by extracting surface meshes from segmented anatomic structure in the SIO dataset.

There are some difficulties to use the SIO data. For example, some errors still present in the surface meshes such as missing triangles or inconsistent normal vector orientation, which would lead to artefacts in the simulated X-ray images. Secondly, both hands are presented in front of the pelvis in the SIO data (Figure 5.2) but only the right hand in the radiograph with AP view (Figure 5.3). The radiographs were taken when the man was alive, the SIO when he was dead. A mismatch of anatomy is therefore to be expected. Lastly, since the data was acquired in the 90s, a film screen might be used to acquire the images rather than a digital equipment and detector. The acquisition parameters used to acquire this radiograph are also unknown and therefore these have been estimated based on those used clinically for the same body parts. Tube voltage, SOD, SDD, cassette orientation, pixel resolution (space between two successive pixels) in the digital image and film sensitivity would have to be guessed.

Our approach is shown in Figure 3.1. To overcome the registration problem, the *source* is the pelvis of the Visible Man model (Figure 5.2 and blue trace in Figure 3.1) and the

target is the pelvic anterior-posterior (AP) radiographic view (Figure 5.3 and red trace in Figure 3.1) The optimisation algorithms will choose different combinations of parameters until a stopping criterion is met (green trace in Figure 3.1). At the end of the registration process, the corresponding simulated images should match the real radiograph.

5.2.3 Multi-objective optimisation

In all registration problems including registration of hands and hips, Non-dominated Sorting Genetic Algorithm III (NSGA-III) is used as the optimisation algorithm. NSGA-III is a well known multi-objective EA. NSGA-III is a many-objective algorithm based on the popular non-dominated sorting approach of NSGA-II [62]. A diversity preservation technique and a survival mechanism are used to ensure the exploration of the Pareto front. Previously, we have demonstrated that NSGA-II has better performance than CMA-ES in terms of performance and speed (Section 4.2). Two metrics (MAE and ZNCC) were used to construct the objective function. Here, we introduce four more popular metrics including structural similarity index measure (SSIM), root mean square error (RMSE), MI and MAPE (further details can be found in Section 2.4). NSGA-III is expected to have better or similar performance comparing to NSGA-II. SSIM measures the structural information between two images [163]. The value ranges between 0 and 1. When SSIM equals to 0, there is no correlation between two images. When SSIM equals to 1, two images are perfectly correlated. RMSE measure differences between two images. Unlike SSIM which focuses on structural information, RMSE consider pixels that are independent to its neighbours. RMSE is always non-negative and best value is 0 where two images are identical. MI is a popular metric in traditional registration problems (see Section 2.2). MI measures the similarity between histograms of two images. MI ranges from 1 (perfectly uncorrelated image values) to 2 (perfectly correlated image values).

How the choice of objective function will affect the result of the registration is investigated here. To align with other metrics which are to be minimised, SSIM, ZNCC and MI are re-written as:

$$\frac{1 - \mathrm{SSIM}(\mathbf{Y}, \hat{\mathbf{Y}})}{2} \tag{5.1}$$

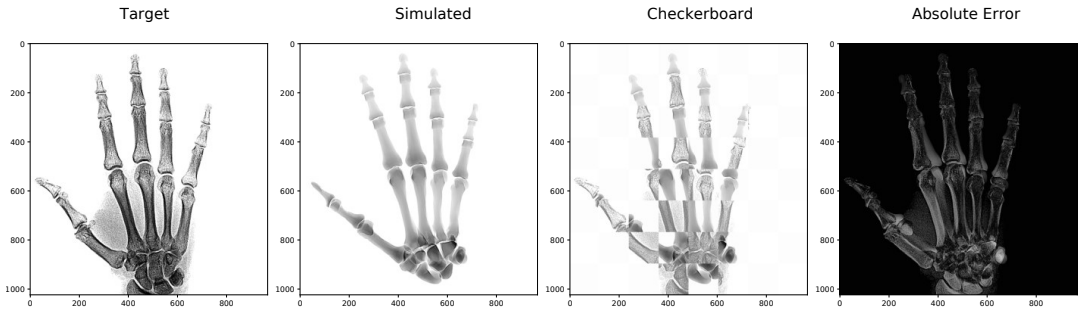


Figure 5.4: *Manual registration of Image 5 by hand after padding, widen boundary ranges and adding a sharpening filter.*

$$\frac{1 - \mathrm{ZNCC}(\mathbf{Y}, \hat{\mathbf{Y}})}{2} \tag{5.2}$$

$$\frac{2 - \mathrm{MI}(\mathbf{Y}, \hat{\mathbf{Y}})}{2} \tag{5.3}$$

All six metrics are used to construct the objective function which is to be minimised by the optimisation algorithm in both registration problems. For hand registration problem, parameters and their corresponding value range would have to be determined after manual registration is conducted. For hip registration problem, there are 8 parameters including rotation of model (2 parameters), the position of the X-ray source (3 parameters) and the position of the detector (3 parameters).

The same stopping criteria are used. The termination tolerance is 1E-5 for both the solution and the function value. Also, the maximum run time for the optimisation process is 20 minutes.

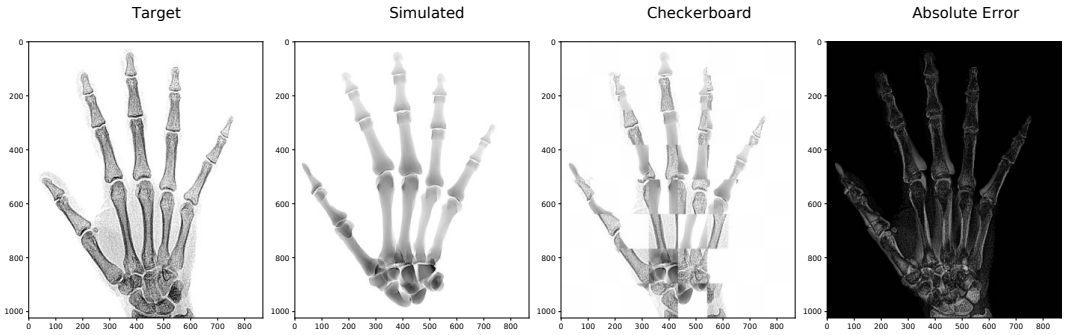


Figure 5.5: *Manual registration of Image 8 by hand after padding, widen boundary ranges and adding a sharpening filter.*

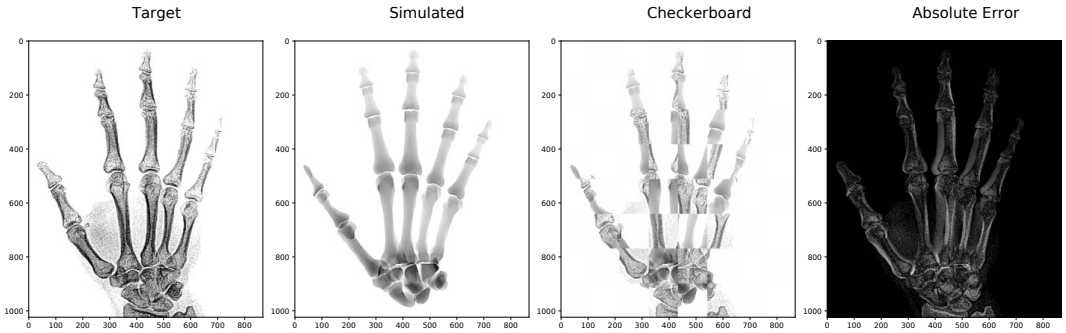


Figure 5.6: *Manual registration of Image 9 by hand after padding, widen boundary ranges and adding a sharpening filter.*

5.3 Results

5.3.1 Hand registration

5.3.1.1 Manual registration of hands

After tuning parameters by hand and utilising the trial-and-error approach, all fingers are able to be recovered in three registration problems. Although image padding was not helping when padding around images, it is still a useful technique. For example, in our case we pad images on the right-hand side to ensure the alignment of hands can be performed at the same position. This is the main reason why different optimisation algorithms tend to have same wrong alignments in previous chapters. Results are shown in Figures 5.4, 5.5 and 5.6. Based on those results, these changes are made in our framework:

Table 5.1: *Rotation and re-scaling parameters to be optimised and their corresponding ranges after manual registration of hands.*

Parameters \ Bones	**Whole hand**	**Thumb**	**Index**	**Middle**	**Fourth**	**Little**
Rotation range (degrees)	[-2, 2]	MC_t: $[-40, 10]$ $[-20, 0]$ PP_t: $[-10, 10]$	MC_i: $[-2, 2]$ $[-2, 2]$ PP_i: $[-30, 10]$ $[-20, 0]$ IP_i: $[-20, 0]$ DP_i: $[-20, 0]$	MC_m: $[-2, 2]$ $[-2, 2]$ PP_m: $[-10, 10]$ $[-20, 0]$ IP_m: $[-20, 0]$ DP_m: $[-20, 0]$	MC_f: $[-2, 2]$ $[-2, 2]$ PP_f: $[-10, 10]$ $[-20, 0]$ IP_f: $[-20, 0]$ DP_f: $[-20, 0]$	MC_l: $[-2, 2]$ $[-2, 2]$ PP_l: $[0, 30]$ $[-20, 0]$ IP_l: $[-20, 0]$ DP_l: $[-20, 0]$
Rescaling ratio	-	$[0.9, 1.1]$	$[0.9, 1.1]$	$[0.9, 1.1]$	$[0.9, 1.1]$	$[0.9, 1.1]$

- **Padding:** 200 pixels are padded on the right hand side for both the target image and predictions in Registration 5, and 100 pixels are padded on the right hand side for both the target image and predictions in Registration 8 and 9.

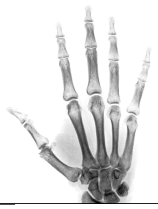
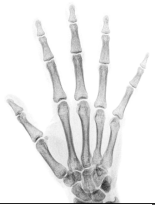
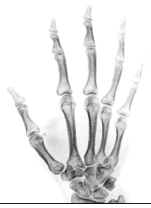
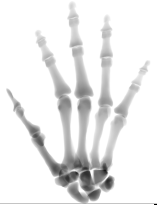
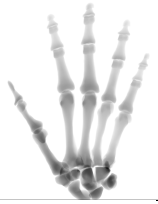
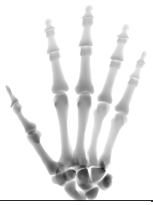
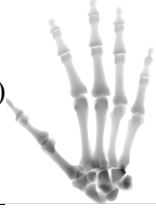
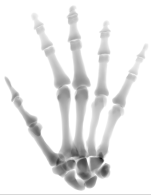
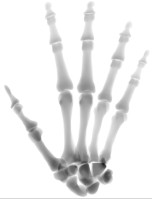
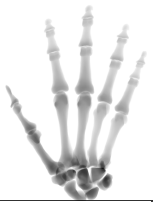
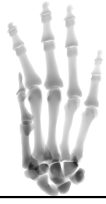
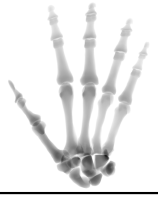
- **Parameters:** The boundary range of MC_t is changed from [-10,10] to [-40,10]. The boundary range of PP_i is changed from [-10,10] to [-30,10]. The boundary range of PP_l is changed from [-10,10] to [0,30]. We also add rotation parameters for bones including MC_i, MC_m, MC_f and MC_l where we aim at better alignment of those bones.

- **Enhance region of interest (ROI):** A sharpening filter is applied on both target image and predicted images to increase the difference between edges of finger bones and finger bones.

The rotation and re-scaling parameters and their corresponding ranges are summarised in Table 5.1. Distance parameters including SOD and SDD are unchanged (details can be found in Section 3.3.3).

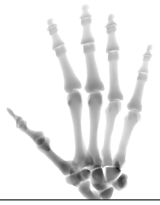
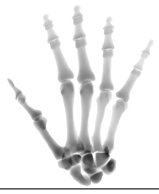
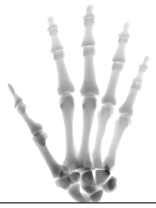
5.3.1.2 Registration of hands

Table 5.2: *Individual solutions with best metric values for Image 5, 8 and 9. Target and predicted images are cropped so that they are comparable with previous results. Same individual solutions for Registration 5 are labelled as: a) the solution with the lowest DZNCC, MAR, RMSE and DMI, b) the solution with lowest MAPE and c) the solution with lowest DSSIM. Corresponding solutions are also labelled in Table 5.3.*

Target and predictions	**Registration 5**	**Registration 8**	**Registration 9**
Target			
Prediction (best MAE)	a) 		
Prediction (best RMSE)	a) 		
Prediction (best MAPE)	b) 		
Prediction (best ZNCC)	a) 		
Prediction (best SSIM)	c) 		

Continued on next page

Table 5.2 – continued from previous page

Target and predictions	Registration 5	Registration 8	Registration 9
Prediction (best MI)	a) 		

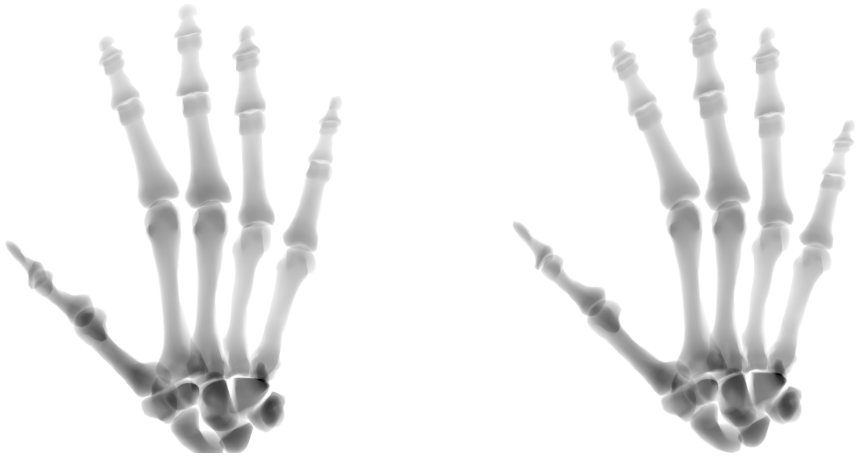
A priori knowledge of why some images are hard to register was obtained by conducting manual registrations. All three images could be registered in the manual registration. Our hypothesis was “if registration can be successfully performed by a human operator, our registration method can achieve similar results and the registration accuracy is expected to be improved". Table 5.2 presents registration results where solutions have extreme metric values in the PF. The second column shows the target and predicted images with best metric value for three registrations. Those solutions with best metric value are obtained from the PF at the last iteration. In Registration 5 the visible improvement was observed in terms of qualities of predicted images except predicted images with best SSIM, comparing with previous results (see Table 4.1). The position of thumb is not able to be recovered with the solution with best MAE. Note those results are solutions with extreme values (best metric values) in the PF. The third and fourth column show the target and predicted images with best metric values for Registration 8 and 9, respectively. All solutions with best metric values produce similar results and has better performance comparing with registration results in Table 4.1. Next, we compare the candidate solutions with best metric value and with median metric values.

The extreme solution is usually not suitable and the ideal solution should be balanced between different metrics, i.e. the solution with median metric values is more appropriate. Results of solutions are displayed in Tables 5.3, 5.4 and 5.5. In Registration 5 two candidate solutions can be selected: one solution with four best metric values (lowest DZNCC, MAE, RMSE and DMI) and one solution with median metric values. Predicted images of two solutions are presented in Figures 5.7a and 5.7b. Both images are similar to each other. In Registration 8

Table 5.3: *Solutions for Registration 5 with six metrics including DZNCC, DSSIM, MAE, RMSE, MAPE and DMI. Corresponding solutions in the second column of Table 5.2 are displayed: a) the solution with the lowest DZNCC, MAR, RMSE and DMI, b) the solution with lowest MAPE and c) the solution with lowest DSSIM. Metrics with lowest value is highlighted in bold.*

Solutions	DZNCC	DSSIM	MAE	RMSE	DMI	MAPE
	0.1916	0.1578	0.4066	0.8754	0.4621	0.0796
	0.1925	0.1569	0.4146	0.8776	0.4622	0.0851
c	0.2037	**0.1548**	0.4205	0.9026	0.4611	0.0855
a	**0.1823**	0.1575	**0.3932**	**0.8539**	**0.4597**	0.0822
b	0.1915	0.1639	0.4056	0.8751	0.4634	**0.0771**

two candidate solutions with median metric values can be selected (Figure 5.8). Solutions are sorted by the value of DZNCC and other metrics can also be used except DSSIM as discussed previously. While two candidate solutions produce similar images for Registration 8, two candidate solutions produce different images for Registration 9. The index finger in one image (Figure 5.9a) is slightly longer than another (Figure 5.9b). Generally speaking, the chosen candidate solutions produce similar images in all three hand registration problems.



(a) **(b)**

Figure 5.7: *Corresponding images of two selected solutions for Registration 5. Solutions are labelled as: a) the solution with four best metric values, and b) the solution with median metric values (see Table 5.3).*

Understanding the relationship between the performance of different metrics in the multi-objective optimisation is important, due to the fact that not all metrics have equal performance. For example, SSIM has the worst performance (visually) among all six metrics in Registration 5. A scatter matrix plot can help us to visualise the relationship between each

Table 5.4: *Solutions for Registration 8 with six metrics including DZNCC, DSSIM, MAE, RMSE, MAPE and DMI. Corresponding images of labelled solutions a) and b) are displayed in Figures 5.8a and 5.8b.*

Solutions	DZNCC	DSSIM	MAE	RMSE	DMI	MAPE
	0.2033	0.1659	0.4195	0.9019	0.4596	0.0854
	0.1962	0.1717	0.4076	0.8859	0.4603	0.0946
	0.1968	0.1677	0.4070	0.8873	0.4581	0.0895
b	0.1967	0.1654	0.4099	0.8869	0.4580	0.0877
	0.1971	0.1664	0.4093	0.8880	0.4580	0.0966
	0.1934	0.1680	0.4032	0.8796	0.4579	0.1029
	0.1925	0.1647	0.4095	0.8775	0.4578	0.0981
	0.2034	0.1695	0.4160	0.9019	0.4602	0.0858
	0.1909	0.1726	0.4040	0.8738	0.4598	0.0921
	0.2009	0.1687	0.4178	0.8964	0.4605	0.0863
	0.1907	0.1707	0.4037	0.8734	0.4591	0.0970
a	0.1965	0.1713	0.4054	0.8865	0.4595	0.0899
	0.1913	0.1690	0.4055	0.8749	0.4586	0.0962
	0.1948	0.1691	0.4047	0.8826	0.4587	0.0993
	0.1971	0.1713	0.4102	0.8879	0.4598	0.0867

metric pairs. We plot the scatter matrices of all six metric pairs for Registrations 5, 8 and 9 in Figures 5.10, 5.11 and 5.12. All individuals in the PF are included and represented by points in these plots. SSIM, ZNCC and MI are in the form of minimisation (Equations 5.1, 5.2 and 5.3) which are presented as DSSIM, DZNCC and DMI. RMSE and DZNCC has the perfect linear relationship, i.e. a straight line can be drew by connect all points in the plot. The value of RMSE and DZNCC increase (or decrease) at the same time. This relationship is observed in all scatter matrix plots for Registration 5, 8 and 9 (subplots of DZNCC-RMSE or RMSE-DZNCC in Figures 5.10, 5.11 and 5.12). The metric pairs of MAE and DZNCC, and MAE and RMSE tend to increase (or decrease) at the same time in Registration 5 (subplots of MAE-RMSE or RMSE-MAE in Figure 5.10) although it is not linear. The relationship becomes less clear in Registration 8 and 9 (Figures 5.11 and 5.12). There seems to be no relationship between SSIM and other metrics, and between DMI and other metrics. MAPE seems to have inverse relationship with other metrics including DZNCC, MAE, RMSE where the value of MAPE increases, the value of DZNCC, MAE, RMSE are in a downward trend.

Table 5.5: *Solutions for Registration 9 with six metrics including DZNCC, DSSIM, MAE, RMSE, MAPE and DMI. Corresponding images of labelled solutions a) and b) are displayed in Figures 5.9a and 5.9b.*

Solutions	DZNCC	DSSIM	MAE	RMSE	DMI	MAPE
	0.2090	0.1767	0.4250	0.9144	0.4654	0.0873
	0.2098	0.1723	0.4213	0.9161	0.4635	0.0880
	0.2070	0.1792	0.4168	0.9098	0.4646	0.0926
	0.2054	0.1760	0.4164	0.9064	0.4639	0.0940
	0.2068	0.1730	0.4149	0.9094	0.4626	0.0981
	0.2047	0.1783	0.4134	0.9048	0.4646	0.0957
a	0.2071	0.1817	0.4202	0.9102	0.4657	0.0898
	0.2128	0.1697	0.4292	0.9227	0.4642	0.0876
	0.2138	0.1725	0.4226	0.9248	0.4633	0.0893
	0.2044	0.1745	0.4131	0.9043	0.4628	0.0974
	0.2081	0.1700	0.4168	0.9123	0.4627	0.0906
b	0.2075	0.1708	0.4091	0.9110	0.4609	0.0922

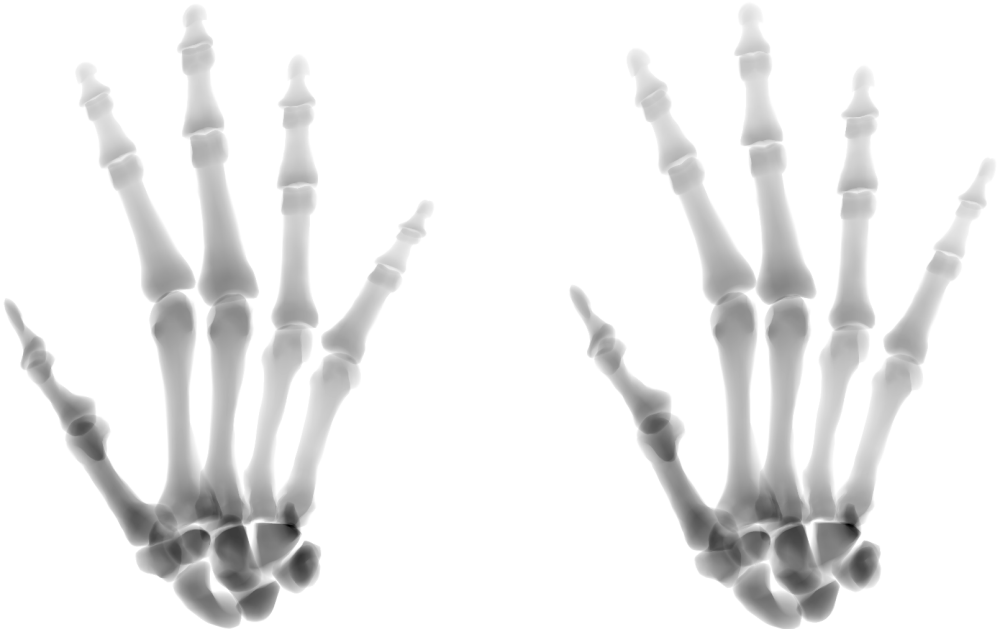
5.3.2 Hip registration

The registration of the VHP male model to a single radiograph with the AP view is more complex than initially expected. The mismatch of anatomy in the VHP male model and the radiograph with AP views would result in pre-exist error in the registration process. Also, the pixel spacing of the film and all parameters that are used to acquire the radiograph are unknown.

Table 5.6: *Solutions in hip registration with six metrics including DZNCC, DSSIM, MAE, RMSE, MAPE and DMI. Images of labelled solutions are displayed in Figure 5.14:a) the solution with four best metric values including DZNCC, DSSIM, MAE and RMSE, and b) and c) are solutions with median metric values sorted using DZNCC.*

Solutions	DZNCC	DSSIM	MAE	RMSE	DMI	MAPE
b	0.2585	0.4333	0.8434	1.0168	0.4709	0.1646
	0.2638	0.3979	0.8371	1.0272	0.4806	0.2192
a	**0.2499**	**0.3792**	**0.8082**	**0.9999**	0.4803	0.2722
c	0.2617	0.4224	0.8247	1.0231	0.4728	0.1899
	0.2582	0.4136	0.8334	1.0162	0.4698	0.1806
	0.2801	0.4302	0.8579	1.0585	0.4722	0.1658
	0.2554	0.4221	0.8251	1.0107	0.4722	0.1826
	0.2854	0.3881	0.8362	1.0685	0.4787	0.2316

Figure 5.13 shows the registration results where DZNCC, DSSIM, MAE and RMSE performs relatively better than DMI and MAPE. Three candidate solutions can be selected (see Table 5.6):



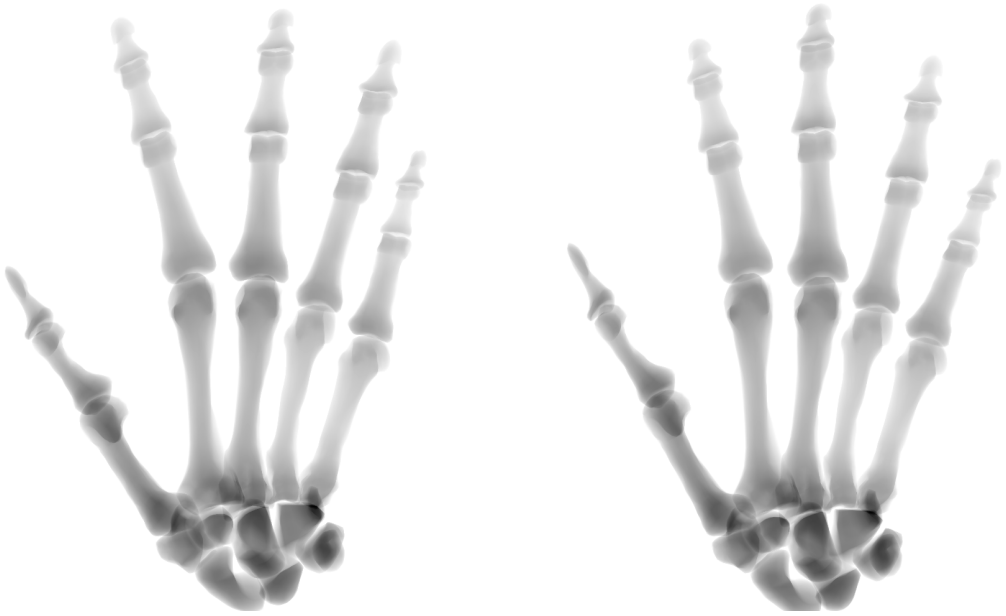
(a) **(b)**

Figure 5.8: *Corresponding images of two selected solutions (median DZNCC) selected in Registration 8 in Table 5.4.*

one solution with four best metric value (best DZNCC, DSSIM, MAE and RMSE) and two solutions with median metric value of DZNCC among its PF. Unlike what we observed in hand registration problems, only one solution with four best metric values can be selected in the hip registration problem. A scatter metric plot of the PF at the last iteration is presented in Figure 5.15. Similar behaviour is observed where DZNCC and RMSE are in a linear relationship where they both increase (or decrease) at the same time. Other metric pairs seem to have little or no relationship.

5.4 Discussion and conclusion

We have demonstrated that our registration method is robust and can generate results accurately with limited computational resources ($\sim$ 20 minutes). We have tested our registration framework on two different problems including registration of hands and hips. Registration of an articulated hand model to the hand radiograph was performed in two steps. Manual registrations of three images were firstly performed to investigate why some images are harder to register than others.



(a) **(b)**

Figure 5.9: *Corresponding images of two selected solutions (median DZNCC) selected in Registration 9 in Table 5.5.*

A priori knowledge was obtained and few techniques were applied to improve our registration framework. Better registration accuracy was achieved using NSGA-III on three test-cases where they were not able to be registered in previous registrations. Registration of hips was performed using NSGA-III where the VHP male model were successfully registered to the radiograph with AP views. This demonstrated the adaptability of our registration method to other registration problems involves different musculoskeletal anatomy.

We have demonstrated that best solutions can be selected in multi-objective optimisation problems. The choice of the best solution varies on two registration problems. In the hand registration problem, solutions with median metric value can be selected as the best solution. However, in the hip registration problem, the best solution is selected has four best metric values. The solutions with median metric value produces unreliable results. One possible reason is that all registrations were performed only once.

The choice of metrics was also investigated. Scatter matrix plots were used to investigate the relationship between different metric pairs. In both hand and hip registrations, we observed that ZNCC has an inverse linear relationship with RMSE, i.e. they are better to be used together than any other metric pairs such as ZNCC and MAPE, or ZNCC and MI. There seems to be little or no relationship between remaining metric pairs. A more thorough analysis of relationships between each metric pair is necessary. One way to achieve that is performing registrations using NSGA-III with different metric pairs. For example, comparing the performance of one objective function constructed by using MAE and ZNCC, with another objective function constructed by using RMSE and ZNCC.

One drawback of our approach was the lack of quantitative analysis. In the multi-objective optimisation problem, there were many solutions with different metric values where they produce similar predicted images. The best solution could not be automatically selected based on metric values.

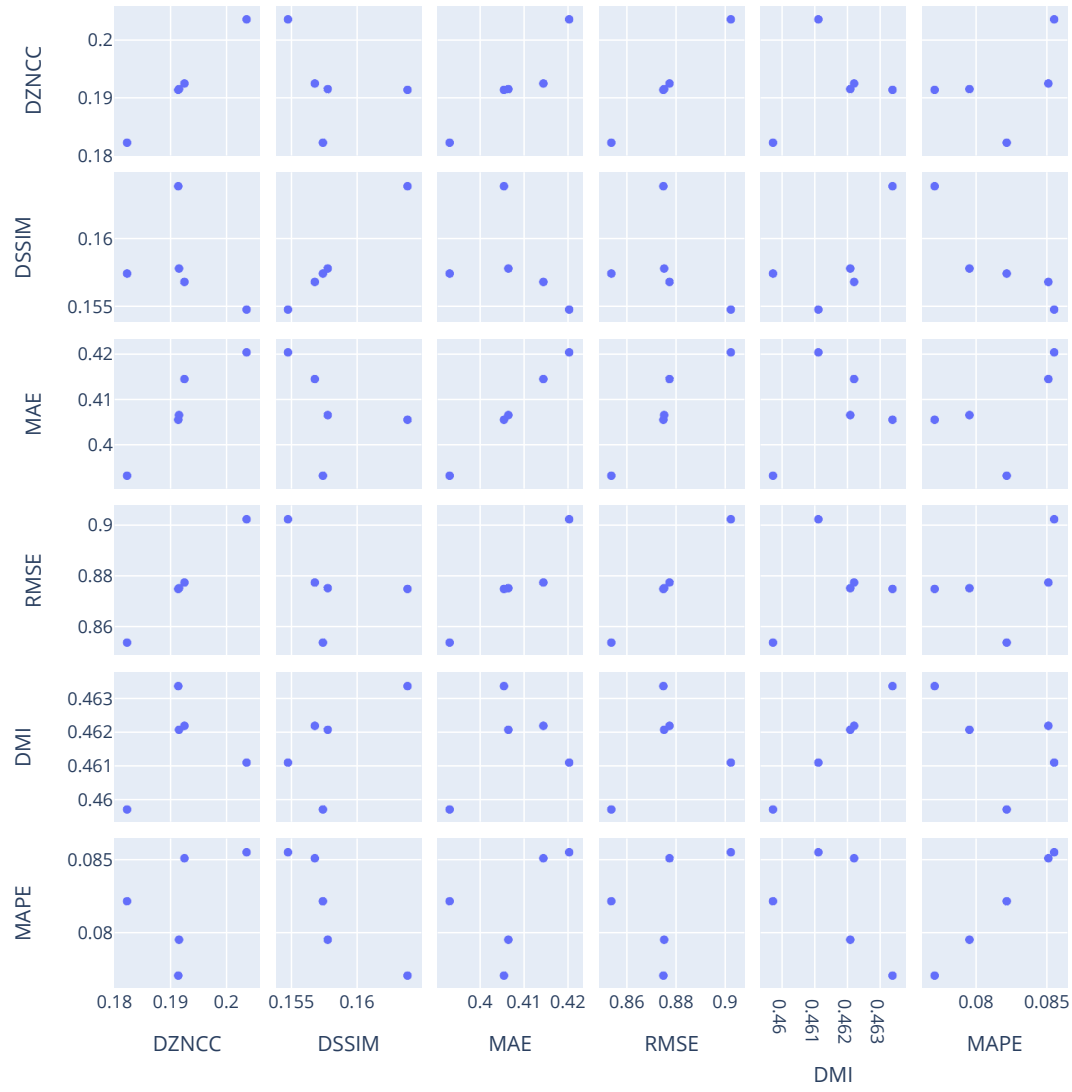


Figure 5.10: *A scatter matrix plot of solutions in a PF for Registration 5. SSIM, ZNCC and MI are in minimisation form (Equations 5.1, 5.2 and 5.3) which are presented as DSSIM, DZNCC and DMI.*

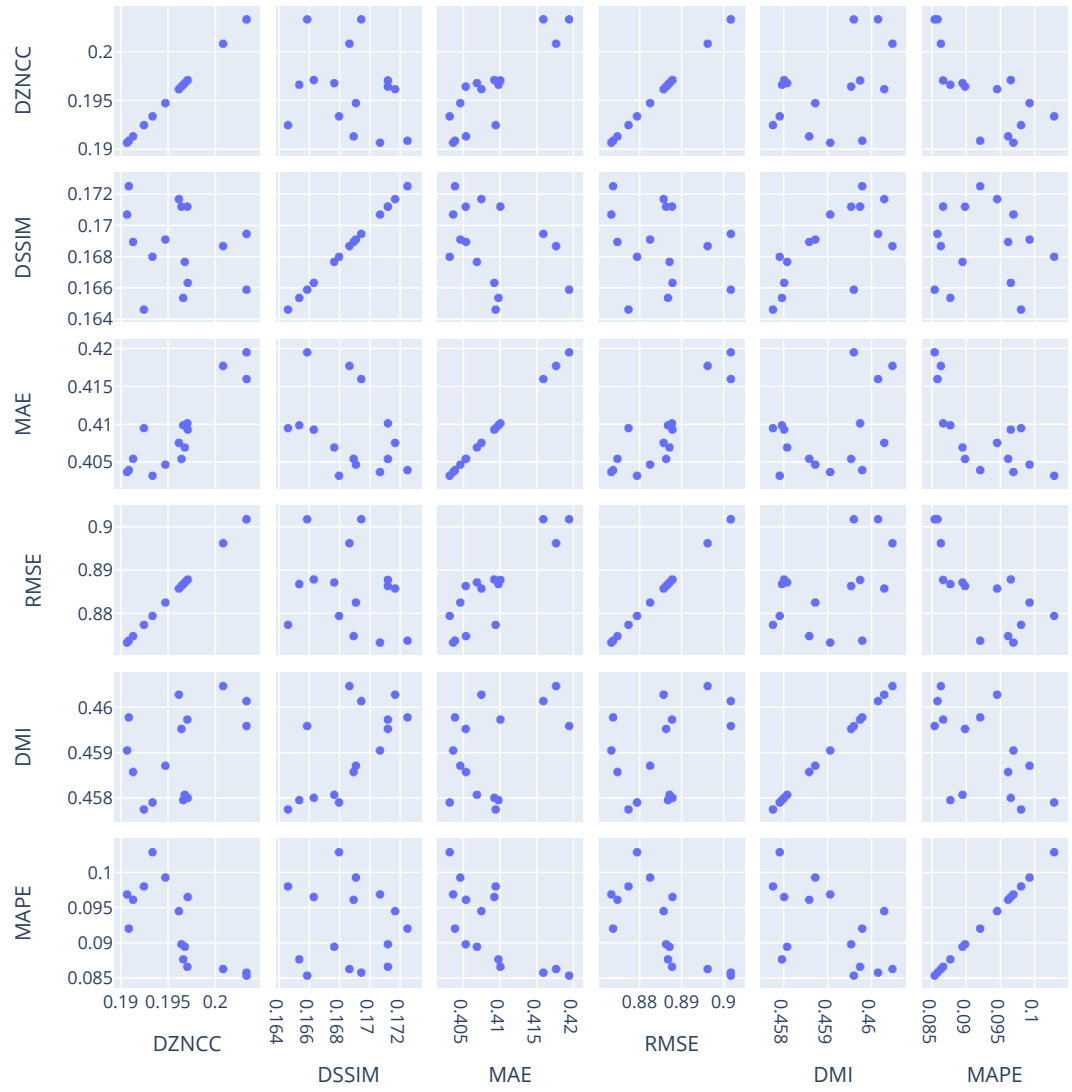


Figure 5.11: *A scatter matrix plot of solutions in a PF for Registration 8. SSIM, ZNCC and MI are in minimisation form (Equations 5.1, 5.2 and 5.3) which are presented as DSSIM, DZNCC and DMI.*

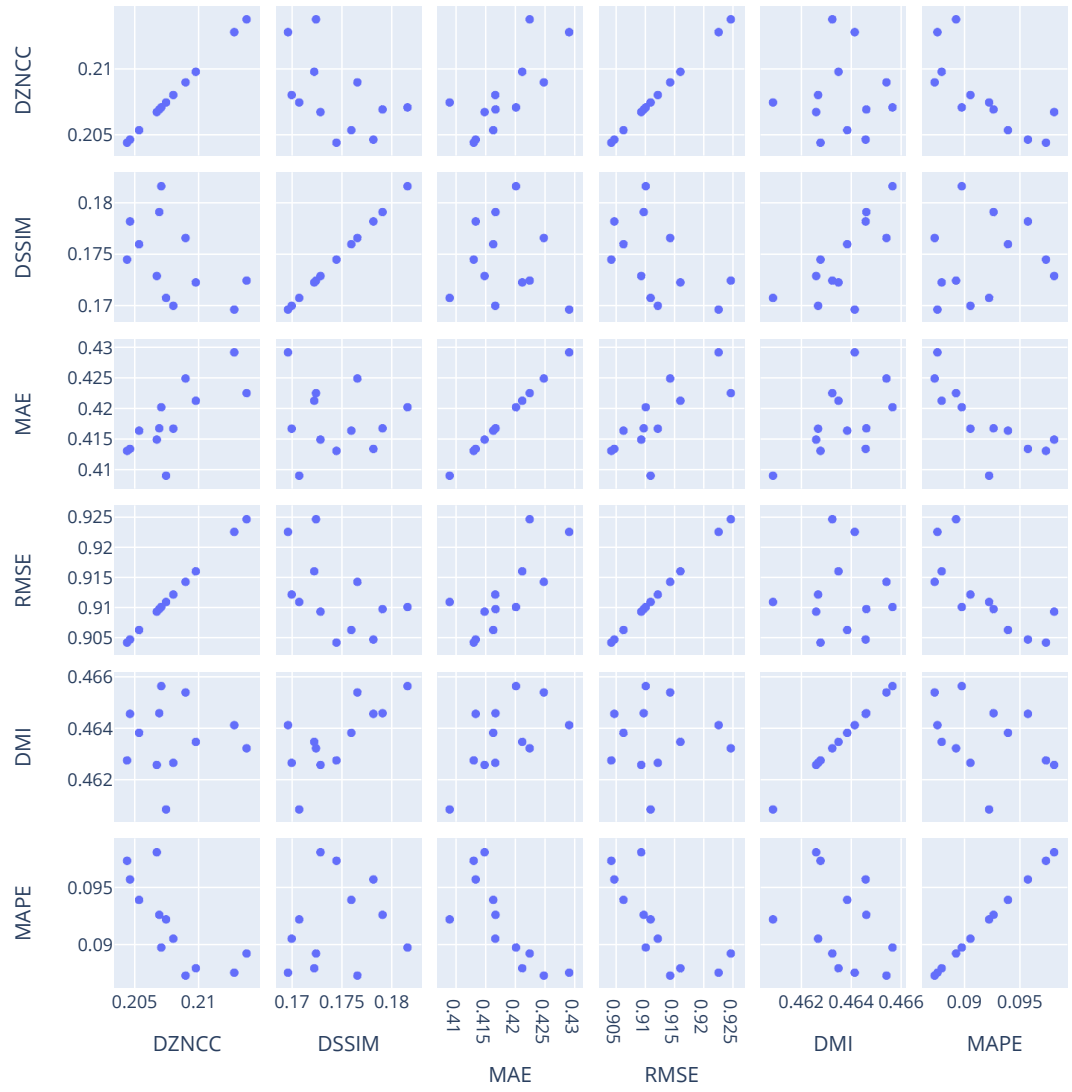


Figure 5.12: *A scatter matrix plot of solutions in a PF for Registration 9. SSIM, ZNCC and MI are in minimisation form (Equations 5.1, 5.2 and 5.3) which are presented as DSSIM, DZNCC and DMI.*

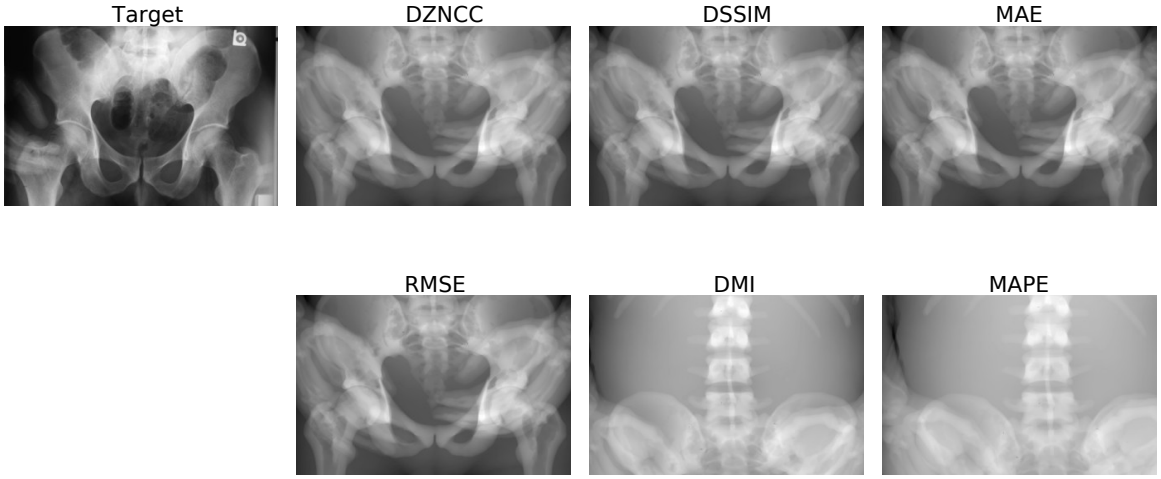
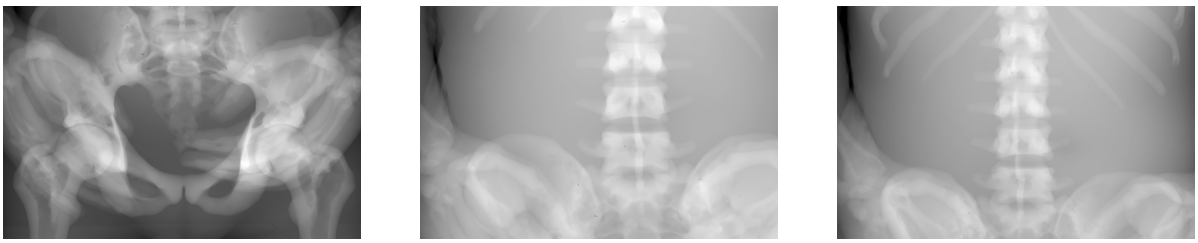


Figure 5.13: *Target and predicted images with best metric values in hip registration with six metrics including DZNCC, DSSIM, MAE, RMSE, DMI and MAPE.*



(a) **(b)** **(c)**

Figure 5.14: *Corresponding images of three selected solutions for the registration of hips. Solutions are sorted based on the value of DZNCC and are labelled as: a) the candidate solution with four best metric value, and b) and c) are two candidate solutions with median metric values (see Table 5.6).*

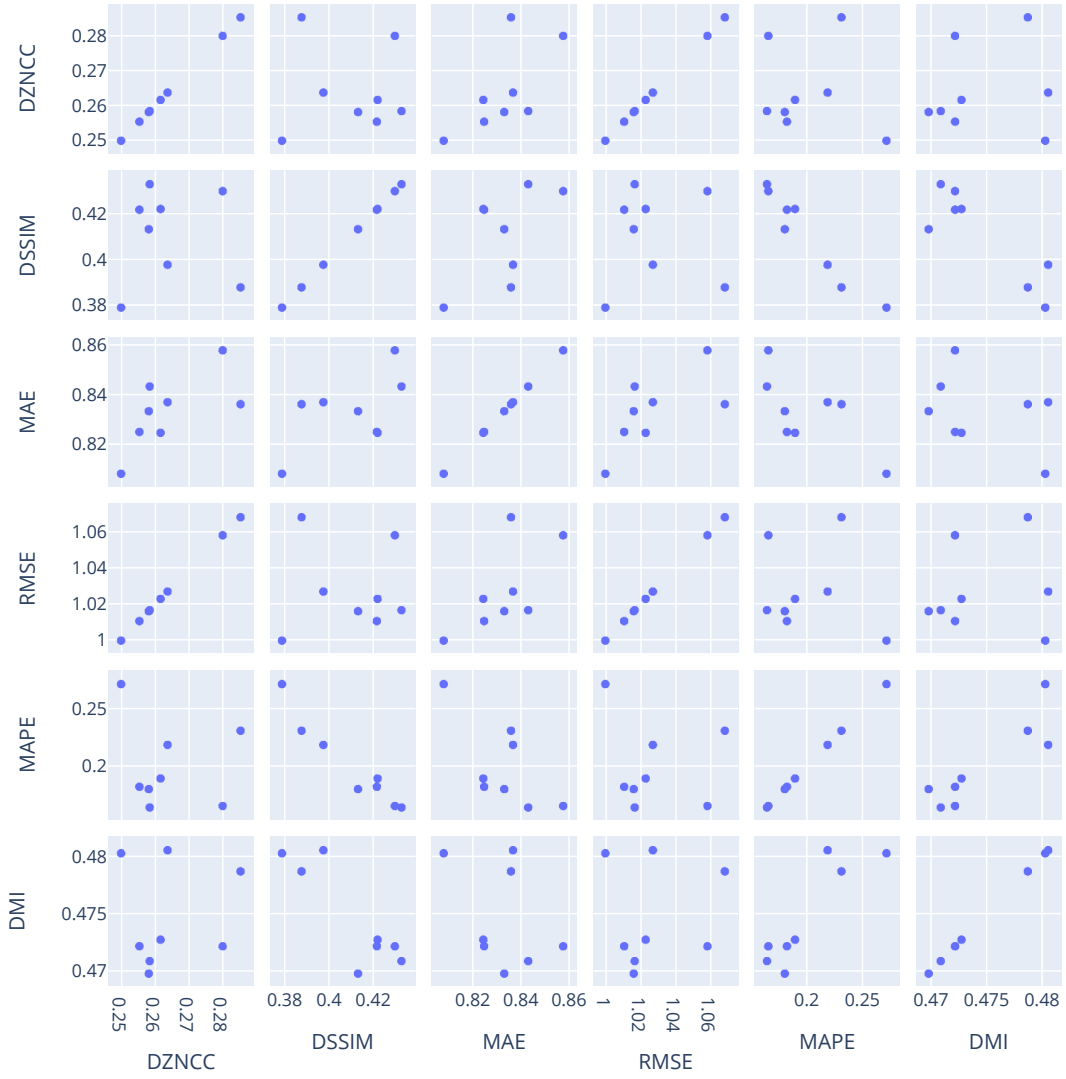


Figure 5.15: *A scatter matrix plot of the PF with different metrics including DZNCC, DSSIM, MAE, RMSE, DMI and MAPE at the last iteration.*

Chapter 6

Discussion and conclusion

6.1 Discussion and limitations

Registration of 3D polygon models to a single image can be achieved using our registration method. Our method relies on fast X-ray simulations, suitable metrics and optimisation algorithms. We have presented both qualitative and quantitative analysis of our method using different datasets. We firstly conducted a proof-of-concept study using two simple test cases and then progressed to registrations of real-world hand radiographs (Chapter 3). Lastly, a lot of effort was devoted into improving our registration method (Chapters 4 and 5) by addressing deficiencies that observed from previous experiments. Different techniques were tested including image padding, sharpening filters and tailored boundary conditions for specific registration problems. Image padding is useful for positioning the hand in the centre of the image. Tailored boundary conditions are proven to be important because limited search space may leads to failure of the whole optimisation process due to the clinical nature of radiographs. Also, we extended our method to another problem, which is the registration of a 3D model to a hip radiograph.

Our registration method can be easily modified to solve any registration problem that requires registration of a 3D polygon model to a single image to be performed. We tested our method on a variety of data such as synthetic hand radiographs, and real clinical radiographs from MURA dataset and VHP dataset. Our registration method performed well in solving real-world problems in terms of both speed and accuracy. However, there are some limitations:

- Our registration method was specifically designed for registration of a 3D model to a single image. Thus, it might not suitable to use in the some cases (e.g. registration of a 3D CT image to X-ray images). Any 3D data must be converted to a 3D model represented by meshes. The development of a 3D model is often time-consuming and requires extensive knowledge in the related research field. Thus, extensive evaluations of our methods in different real-world problems were not possible due to the lack of datasets. One solution is creating a dataset that contains 3D models and related radiographs (e.g. A 3D model of the whole human body and the related radiographs).

- There was no skins presented in our 3D hand model. Normally, hand radiographs contains both bones and skins. We removed skins as much as possible without affecting bones in the real hand radiograph. However, there were still some skins left in the real hand radiographs which resulted in pre-existing errors before registration. There are two ways to address this problem. The first one is adding a skin model to the existing hand model and skins are kept in the hand radiographs. However, by adding skins to both real and simulated hand radiographs, the complexity of the problem increases since the optimisation algorithms will have to match both skins and bones. Also, we have to taken into account the development of a skin model itself. An alternative way is automatically removing the skins using segmentation techniques. In this way the optimisation algorithms would only need to match bones in both real and simulated hand radiographs. However, segmentation might produces errors. For example, a small section of bones might be removed along with skins.

- In our registration method only a single radiograph with PA view is used. This might cause a problem which is known as depth ambiguity where registration of the same radiograph leads to recovered 3D hands with different sizes (depth of bones). One way to fix this problem is to register the same radiograph but with the lateral view (or ball-catching poses) so that the missing depth information could be obtained.

- Bi- and multi-objective optimisation algorithms generally performed better than single-objective optimisation algorithms. The best solution from a single-objective optimisation can be easily identified. However, this was not the case for bi- and multi-objective optimisations where a list of candidate solutions (PF) was available. We considered two cases to select one best solution. The first one is relatively easy. If selected solutions provide different results, then the best solution is the one produces an image that best matches the target image. The second one is, if selected solutions provided similar results, the solution with median metric values is chosen as the best solution. In both cases manual comparison is required to identify the best solution such as the use of error maps or other image comparison techniques. Human interventions might not be possible if thousands or more registration processes are performed, which prevents a thorough evaluation of our method to be conducted, i.e. testing our methods using thousands of available images to gain statistically reliable analysis is not possible. Also, all registrations need be repeated for several times (e.g. 15 times) to gather statistically meaningful results.

6.2 Conclusion and future work

In this thesis we reviewed the concept of registration and 3D/2D registration. Also, related 3D/2D registration methods, metrics and optimisation algorithms were presented and discussed. We identified that the registration of polygon meshes has been overlooked and we proposed to address this. We proposed and evaluated our 3D/2D registration method that aims to register a 3D polygon model to a single image using a variety of datasets and different metrics and optimisation algorithms. Our method is found to be robust and produces good results. However, there are some drawbacks of our method as discussed in the previous section. We plan to address those problems in the future and provide an insight of further researches:

- Further evaluation of our registration method (see Chapter 5) on all 15 images and repeating the registration for 15 times for each image. Comparisons between the results with six metrics and with different metric pairs are necessary, i.e. comparing the

performance of registration using 6 metrics with 2 metrics. In this way best metric pairs might be used in further experiments and reduces time required to do the optimisation. Total number of registrations when 6 metrics are used is 15 and is 225 when different metric pairs are used. By repeating all registrations for 15 times, the total number of registrations becomes 3600. It will take roughly 50 days using our computer where hardware specifications were presented in Section 3.3.4.1. More powerful computers (e.g. Supercomputing Wales, available at `https://www.supercomputing.wales/`) is required to shorten the registration process.

- Making use of image segmentation techniques to pre-process the real hand radiographs automatically (e.g. iterative thresholding [162]) and developing a visualisation tool using graphing libraries (e.g. Plotly, available at `https://github.com/plotly`) which could speed up the selection process in the multi-objective optimisation.

- We could create a dataset that contains 3D mesh models and their corresponding radiograph with PA view or radiographs with multiple views. There are available datasets containing 3D volumetric data (e.g. *Mosmed-1110* including 1110 3D CT scans of patients [97]). The 3D mesh models can be created from those data using deep learning techniques (e.g. *Voxel2Mesh* [168]). Also, fast DRR generations from 3D volumes can be achieved [149]. This dataset will be very useful for testing, evaluating and improving our method and for evaluating the applicability of our approach to other articulated musculoskeletal anatomy.

Appendix A

List of publications

- Wen, T., Mihail, R., Al-Maliki, S.F., Létang, J.M. and Vidal, F.P., 2019. *Registration of 3D Triangular Models to 2D X-ray Projections Using Black-box Optimisation and X-ray Simulation*. In CGVC (pp. 105-113).

- Wen, T., Mihail, R. and Vidal, F., *3D-2D Registration Using X-Ray Simulation and CMA-ES*. In Applications of Evolutionary Computation: 24th International Conference, EvoApplications 2021, April 7–9, 2021, Proceedings (Vol. 12694, p. 453). Springer Nature.

- Pointon, J.L., Wen, T., Tugwell-Allsup, J., Sújar, A., Létang, J.M. and Vidal, F.P., 2023. *Simulation of X-ray projections on GPU: Benchmarking gVirtualXray with clinically realistic phantoms.* Computer Methods and Programs in Biomedicine, 234, p.107500.

- Pointon, J.L., Wen, T., Tugwell-Allsup, J., Létang, J.M. and Vidal, F.P., 2023. *gVirtualXray (gVXR): Simulating X-ray radiographs and CT volumes of anthropomorphic phantoms*. Software Impacts, 16, p.100513.

References

[1] M. Ackerman, 'The visible human project,' *Proceedings of the IEEE*, vol. 86, no. 3, pp. 504–511, 1998. DOI: 10.1109/5.662875 (pp. 5, 8, 89).

[2] S. Agostinelli *et al.*, 'Geant4–a simulation toolkit,' *Nucl Instrum Methods Phys Res A*, vol. 506, no. 3, pp. 250–303, 2003, ISSN: 0168-9002. DOI: 10.1016/S0168-9002(03)01368-8 (p. 40).

[3] S. Agostinelli, J. Allison, K. a. Amako *et al.*, 'Geant4—a simulation toolkit,' *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003 (p. 43).

[4] S. Aouadi and L. Sarry, 'Accurate and precise 2D–3D registration based on x-ray intensity,' *Computer vision and image understanding*, vol. 110, no. 1, pp. 134–151, 2008 (pp. 20, 21, 68).

[5] S. Arora, J. Acharya, A. Verma and P. K. Panigrahi, 'Multilevel thresholding for image segmentation through a fast statistical recursive algorithm,' *Pattern Recognition Letters*, vol. 29, no. 2, pp. 119–125, 2008 (p. 15).

[6] J.-F. Aubry, L. Beaulieu, L.-M. Girouard *et al.*, 'Measurements of intrafraction motion and interfraction and intrafraction rotation of prostate by three-dimensional analysis of daily portal imaging with radiopaque markers,' *International Journal of Radiation Oncology* Biology* Physics*, vol. 60, no. 1, pp. 30–39, 2004 (pp. 12, 13).

[7] B. B. Avants, C. L. Epstein, M. Grossman and J. C. Gee, 'Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain,' *Medical image analysis*, vol. 12, no. 1, pp. 26–41, 2008 (p. 23).

[8] T. Bäck and H. Schwefel, 'An overview of evolutionary algorithms for parameter optimization,' *Evolutionary Computation*, vol. 1, no. 1, pp. 1–23, 1993, ISSN: 1063-6560. DOI: 10.1162/evco.1993.1.1.1 (p. 35).

[9] J. Baró, J. Sempau, J. M. Fernández-Varea and F. Salvat, 'PENELOPE: An algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter,' *Nucl Instrum Methods Phys Res B*, vol. 100, no. 1, pp. 31–46, 1995. DOI: 10.1016/0168-583X(95)00349-5 (p. 40).

[10] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli and J. de Guise, '3d/2d registration and segmentation of scoliotic vertebrae using statistical models,' *Computerized Medical Imaging and Graphics*, vol. 27, no. 5, pp. 321–337, 2003 (p. 16).

[11] S. S. M. e. a. BERGER M. J. HUBBELL J. H., *XCOM: Photon cross section database*, 2010. [Online]. Available: https://www.nist.gov/pml/xcom-photon-cross-sections-database (p. 43).

[12] J. Bert *et al.*, 'Hybrid GATE: A GPU/CPU implementation for imaging and therapy applications,' in *2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC)*, Oct. 2012, pp. 2247–2250. DOI: 10.1109/NSSMIC.2012.6551511 (p. 43).

[13] A. Bielajew *et al.*, 'History, overview and recent improvements of EGS4,' Stanford Linear Accelerator Center (SLAC), Tech. Rep., 2004. DOI: 10.2172/839781 (p. 40).

[14] J. Bijhold, 'Three-dimensional verification of patient placement during radiotherapy using portal images,' *Medical physics*, vol. 20, no. 2, pp. 347–356, 1993 (p. 16).

[15] W. Birkfellner, M. Stock, M. Figl *et al.*, 'Stochastic rank correlation: A robust merit function for 2D/3D registration of image data obtained at different energies,' *Medical physics*, vol. 36, no. 8, pp. 3420–3428, 2009 (pp. 21, 24, 25, 30).

[16] W. Birkfellner, J. Wirth, W. Burgstaller *et al.*, 'A faster method for 3D/2D medical image registration—a simulation study,' *Physics in Medicine & Biology*, vol. 48, no. 16, p. 2665, 2003 (pp. 21, 25).

[17] K. K. Brock, 'Imaging and image-guided radiation therapy in liver cancer,' in *Seminars in radiation oncology*, Elsevier, vol. 21, 2011, pp. 247–255 (p. 11).

[18] L. G. Brown and T. E. Boult, 'Registration of planar film radiographs with computed tomography,' in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, IEEE, 1996, pp. 42–51 (pp. 21, 25).

[19] J. Byrne, C. Colominas, J. Hipwell *et al.*, 'Assessment of a technique for 2D–3D registration of cerebral intra-arterial angiography,' *The British journal of radiology*, vol. 77, no. 914, pp. 123–128, 2004 (p. 25).

[20] X. Cao, J. Yang, J. Zhang *et al.*, 'Deformable image registration based on similarity-steered cnn regression,' in *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*, Springer, 2017, pp. 300–308 (pp. 22–24).

[21] A. G. Chandler, R. J. Pinder, T. Netsch *et al.*, 'Correction of misaligned slices in multi-slice cardiovascular magnetic resonance using slice-to-volume registration,' *Journal of cardiovascular magnetic resonance*, vol. 10, no. 1, pp. 1–9, 2008 (p. 11).

[22] S. D. Chang, W. Main, D. P. Martin, I. C. Gibbs and M. P. Heilbrun, 'An analysis of the accuracy of the cyberknife: A robotic frameless stereotactic radiosurgical system,' *Neurosurgery*, vol. 52, no. 1, pp. 140–147, 2003 (p. 11).

[23] J.-H. Choi, D.-W. Seo, D. H. Park, S. K. Lee and M.-H. Kim, 'Fiducial placement for stereotactic body radiation therapy under only endoscopic ultrasonography guidance in pancreatic and hepatic malignancy: Practical feasibility and safety,' *Gut and Liver*, vol. 8, no. 1, p. 88, 2014 (p. 16).

[24] C.-R. Chou and S. Pizer, 'Local regression learning via forest classification for 2d/3d deformable registration,' in *Medical Computer Vision. Large Data in Medical Imaging: Third International MICCAI Workshop, MCV 2013, Nagoya, Japan, September 26, 2013, Revised Selected Papers 3*, Springer, 2014, pp. 24–33 (p. 11).

[25] S. ébastien Clippe, D. Sarrut, C. Malet, S. Miguet, C. Ginestet and C. Carrie, 'Patient setup error measurement using 3D intensity-based image registration techniques,' *International Journal of Radiation Oncology* Biology* Physics*, vol. 56, no. 1, pp. 259–265, 2003 (pp. 21, 25, 29, 30).

[26] G. B. Coleman and H. C. Andrews, 'Image segmentation by clustering,' *Proceedings of the IEEE*, vol. 67, no. 5, pp. 773–785, 1979 (p. 15).

[27] P. Collet, E. Lutton, F. Raynal and M. Schoenauer, 'Polar IFS+Parisian Genetic Programming=Efficient IFS Inverse Problem Solving,' *Genetic Programming and Evolvable Machines*, vol. 1, no. 4, pp. 339–361, Oct. 2000, ISSN: 1573-7632. DOI: 10.1023/A:1010065123132 (p. 36).

[28] O. Dandekar, W. Plishker, S. S. Bhattacharyya and R. Shekhar, 'Multiobjective optimization for reconfigurable implementation of medical image registration,' *International Journal of Reconfigurable Computing*, vol. 2008, 2008 (p. 69).

[29] C. Darwin, *On the origin of species by means of natural selection*. John Murray (London), 1859 (p. 35).

[30] P. De Bruin, B. Kaptein, B. Stoel, J. Reiber, P. Rozing and E. Valstar, 'Image-based rsa: Roentgen stereophotogrammetric analysis based on 2D-3D image registration,' *Journal of biomechanics*, vol. 41, no. 1, pp. 155–164, 2008 (pp. 21, 25).

[31] J. De Los Santos, R. Popple, N. Agazaryan *et al.*, 'Image guided radiation therapy (igrt) technologies for radiation therapy localization and delivery,' *International journal of radiation oncology, biology, physics*, vol. 87, no. 1, pp. 33–45, 2013 (p. 11).

[32] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, 'A fast and elitist multiobjective genetic algorithm: Nsga-ii,' *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002 (pp. 38, 69).

[33] J. Dey and S. Napel, 'Targeted 2D/3D registration using ray normalization and a hybrid optimizer,' *Medical physics*, vol. 33, no. 12, pp. 4730–4738, 2006 (pp. 21, 25).

[34] L. Di Stefano, S. Mattoccia and F. Tombari, 'Zncc-based template matching using bounded partial correlation,' *Pattern recognition letters*, vol. 26, no. 14, pp. 2129–2134, 2005 (p. 25).

[35] A. Elen, J. Hermans, J. Ganame *et al.*, 'Automatic 3-d breath-hold related motion correction of dynamic multislice mri,' *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 868–878, 2010 (p. 11).

[36] B. Fei, J. L. Duerk, D. T. Boll, J. S. Lewin and D. L. Wilson, 'Slice-to-volume registration and its potential application to interventional mri-guided radio-frequency thermal ablation of prostate cancer,' *IEEE transactions on medical imaging*, vol. 22, no. 4, pp. 515–525, 2003 (p. 11).

[37] J. Feldmar, N. Ayache and F. Betting, '3d–2d projective registration of free-form curves and surfaces,' *Computer vision and image understanding*, vol. 65, no. 3, pp. 403–424, 1997 (p. 16).

[38] C. M. Fonseca, L. Paquete and M. López-Ibánez, 'An improved dimension-sweep algorithm for the hypervolume indicator,' in *2006 IEEE international conference on evolutionary computation*, IEEE, 2006, pp. 1157–1163 (p. 70).

[39] M. D. Foote, B. E. Zimmerman, A. Sawant and S. C. Joshi, 'Real-time 2d-3d deformable registration with deep learning and application to lung radiotherapy targeting,' in *Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, Proceedings 26*, Springer, 2019, pp. 265–276 (p. 11).

[40] N. Freud, P. Duvauchelle, J. M. Létang and D. Babot, 'Fast and robust ray casting algorithms for virtual X-ray imaging,' *Nucl Instrum Methods Phys Res B*, vol. 248, no. 1, pp. 175–180, 2006, ISSN: 0168-583X. DOI: 10.1016/j.nimb.2006.03.009 (pp. 40, 43).

[41] A. H. Friedman, 'Advanced techniques in image-guided brain and spine surgery.,' *Neuro-Oncology*, vol. 6, no. 1, p. 77, 2004 (p. 11).

[42] D. Fu and G. Kuduvalli, 'A fast, accurate, and automatic 2D-3D image registration for image-guided cranial radiosurgery,' *Medical physics*, vol. 35, no. 5, pp. 2180–2194, 2008 (pp. 21, 25, 28).

[43] A. Gardi, R. Sabatini and S. Ramasamy, 'Multi-objective optimisation of aircraft flight trajectories in the atm and avionics context,' *Progress in Aerospace Sciences*, vol. 83, pp. 1–36, 2016 (p. 69).

[44] O. Ghasemalizadeh, S. Khaleghian and S. Taheri, 'A review of optimization techniques in artificial networks,' *Int. J. Adv. Res*, vol. 4, no. 9, pp. 1668–1686, 2016 (p. 35).

[45] A. Ghosh, N. R. Pal and S. K. Pal, 'Image segmentation using a neural network,' *Biological Cybernetics*, vol. 66, no. 2, pp. 151–158, 1991 (p. 15).

[46] F. Godenschweger, U. Kägebein, D. Stucht *et al.*, 'Motion correction in mri of the brain,' *Physics in medicine & biology*, vol. 61, no. 5, R32, 2016 (p. 11).

[47] R. H. Gong, A. J. Stewart and P. Abolmaesumi, 'A new method for ct to fluoroscope registration based on unscented kalman filter,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2006, pp. 891–898 (pp. 18, 21, 25).

[48] R. N. Graham, R. W. Perriss and A. F. Scarsbrook, 'DICOM demystified: A review of digital file formats and their use in radiological practice,' *Clinical radiology*, vol. 60, no. 11, pp. 1133–1140, 2005 (p. 52).

[49] C. Gray, S. Al-Maliki and F. Vidal, 'Data exploration in evolutionary reconstruction of pet images,' English, *Genetic Programming and Evolvable Machines*, vol. 19, no. 3, pp. 391–419, Sep. 2018, ISSN: 1389-2576. DOI: 10.1007/s10710-018-9330-7 (p. 37).

[50] M. Groher, T. F. Jakobs, N. Padoy and N. Navab, 'Planning and intraoperative visualization of liver catheterizations: New cta protocol and 2d-3d registration method,' *Academic radiology*, vol. 14, no. 11, pp. 1325–1340, 2007 (p. 16).

[51] N. Hansen and A. Ostermeier, 'Completely derandomized self-adaptation in evolution strategies,' *Evol Comput*, vol. 9, no. 2, pp. 159–195, 2001 (pp. 38, 53).

[52] R. M. Haralick and L. G. Shapiro, 'Image segmentation techniques,' *Computer vision, graphics, and image processing*, vol. 29, no. 1, pp. 100–132, 1985 (p. 14).

[53] G. Hinton, L. Deng, D. Yu *et al.*, 'Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,' *IEEE Signal processing magazine*, vol. 29, no. 6, pp. 82–97, 2012 (p. 10).

[54] J. H. Hipwell, G. P. Penney, R. A. McLaughlin *et al.*, 'Intensity-based 2-D-3-D registration of cerebral angiograms,' *IEEE transactions on medical imaging*, vol. 22, no. 11, pp. 1417–1426, 2003 (pp. 21, 25).

[55] A. K. Ho, D. Fu, C. Cotrutz *et al.*, 'A study of the accuracy of cyberknife spinal radiosurgery using skeletal structure tracking,' *Operative Neurosurgery*, vol. 60, no. suppl_2, ONS–147, 2007 (p. 25).

[56] K. H. Höhne, B. Pflesser, A. Pommert *et al.*, 'A realistic model of human structure from the visible human data,' *Methods of information in medicine*, vol. 40, no. 02, pp. 83–89, 2001 (p. 90).

[57] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, 'Densely connected convolutional networks,' in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708 (p. 11).

[58] A. Hurvitz and L. Joskowicz, 'Registration of a ct-like atlas to fluoroscopic x-ray images using intensity correspondences,' *International journal of computer assisted radiology and surgery*, vol. 3, no. 6, p. 493, 2008 (pp. 20, 21, 68).

[59] H. Imamura, N. Ida, N. Sugimoto *et al.*, 'Registration of preoperative cta and intraoperative fluoroscopic images for assisting aortic stent grafting,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2002, pp. 477–484 (pp. 21, 25).

[60] A. Jacob, P. Regazzoni, D. Bilecen, M. Rasmus, R. Huegli and P. Messmer, 'Medical technology integration: Ct, angiography, imaging-capable or-table, navigation and robotics in a multifunctional sterile suite,' *Minimally Invasive Therapy & Allied Technologies*, vol. 16, no. 4, pp. 205–211, 2007 (p. 21).

[61] D. Jaffray, P. Kupelian, T. Djemil and R. M. Macklis, 'Review of image-guided radiation therapy,' *Expert review of anticancer therapy*, vol. 7, no. 1, pp. 89–103, 2007 (p. 11).

[62] H. Jain and K. Deb, 'An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, Part II: Handling constraints and extending to an adaptive approach,' *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 602–622, 2014. DOI: 10.1109/TEVC.2013.2281534 (p. 91).

[63] H.-S. Jans, A. Syme, S. Rathee and B. Fallone, '3D interfractional patient position verification using 2D-3D registration of orthogonal images,' *Medical physics*, vol. 33, no. 5, pp. 1420–1439, 2006 (p. 25).

[64] X. Jia, P. Ziegenhein and S. B. Jiang, 'GPU-based high-performance computing for radiation therapy,' *Phys Med Biol*, vol. 59, no. 4, R151–R182, 2014. DOI: 10.1088/0031-9155/59/4/R151 (p. 43).

[65] J.-Y. Jin, S. Ryu, K. Faber *et al.*, '2d/3d image fusion for accurate target localization and evaluation of a mask based stereotactic system in fractionated stereotactic radiotherapy of cranial lesions,' *Medical physics*, vol. 33, no. 12, pp. 4557–4566, 2006 (p. 12).

[66] J. Jin, C. Yang and Y. Zhang, 'An improved cma-es for solving large scale optimization problem,' in *Advances in Swarm Intelligence: 11th International Conference, ICSI 2020, Belgrade, Serbia, July 14–20, 2020, Proceedings 11*, Springer, 2020, pp. 386–396 (p. 38).

[67] E. Jones, T. Oliphant, P. Peterson *et al.*, *SciPy: Open source scientific tools for Python*, [Online; accessed 08/08/2019], 2001. [Online]. Available: `http://www.scipy.org/` (p. 49).

[68] S. Jonic, P. Thévenaz and M. A. Unser, 'Multiresolution-based registration of a volume to a set of its projections,' in *Medical Imaging 2003: Image Processing*, International Society for Optics and Photonics, vol. 5032, 2003, pp. 1049–1052 (pp. 21, 25).

[69] N. Katoh, H. Onishi, Y. Uchinami *et al.*, 'Real-time tumor-tracking radiotherapy and general stereotactic body radiotherapy for adrenal metastasis in patients with oligometastasis,' *Technology in cancer research & treatment*, vol. 17, p. 1 533 033 818 809 983, 2018 (p. 12).

[70] E. Kerrien, M.-O. Berger, E. Maurincomme, L. Launay, R. Vaillant and L. Picard, 'Fully automatic 3D/2D subtracted angiography registration,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 1999, pp. 664–671 (pp. 21, 25).

[71] A. Khamene, P. Bloch, W. Wein, M. Svatos and F. Sauer, 'Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy,' *Medical image analysis*, vol. 10, no. 1, pp. 96–112, 2006 (p. 25).

[72] B. Kim, J. L. Boes, P. H. Bland, T. L. Chenevert and C. R. Meyer, 'Motion correction in fmri via registration of individual slices into an anatomical volume,' *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 41, no. 5, pp. 964–972, 1999 (p. 11).

[73] J. Kim, J. A. Fessler, K. L. Lam, J. M. Balter and R. K. Ten Haken, 'A feasibility study of mutual information based setup error estimation for radiotherapy,' *Medical Physics*, vol. 28, no. 12, pp. 2507–2517, 2001 (pp. 21, 25).

[74] J. Kim, S. Li, D. Pradhan *et al.*, 'Comparison of similarity measures for rigid-body ct/dual x-ray image registrations,' *Technology in cancer research & treatment*, vol. 6, no. 4, pp. 337–345, 2007 (pp. 21, 25, 28).

[75] J. Kim, F.-F. Yin, Y. Zhao and J. H. Kim, 'Effects of x-ray and CT image enhancements on the robustness and accuracy of a rigid 3D/2D image registration,' *Medical physics*, vol. 32, no. 4, pp. 866–873, 2005 (pp. 21, 25).

[76] Y. Kita, D. L. Wilson and J. A. Noble, 'Real-time registration of 3d cerebral vessels to x-ray angiograms,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 1998, pp. 1125–1133 (p. 15).

[77] D. Knaan and L. Joskowicz, 'Effective intensity-based 2D/3D rigid registration between fluoroscopic x-ray and ct,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2003, pp. 351–358 (pp. 21, 25).

[78] W. Krattenthaler, K. J. Mayer and M. Zeiller, 'Point correlation: A reduced-cost template matching technique,' in *Proceedings of 1st International Conference on Image Processing*, IEEE, vol. 1, 1994, pp. 208–212 (p. 25).

[79] A. Krizhevsky, I. Sutskever and G. E. Hinton, 'Imagenet classification with deep convolutional neural networks,' *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017 (p. 10).

[80] D. LaRose, J. Bayouth and T. Kanade, 'Transgraph: Interactive intensity-based 2D/3D registration of x-ray and CT data,' in *Medical Imaging 2000: Image Processing*, International Society for Optics and Photonics, vol. 3979, 2000, pp. 385–396 (pp. 21, 25).

[81] D. LaRose, L. Cassenti, B. Jaramaz, J. Moody, T. Kanade and A. DiGioia, 'Post-operative measurement of acetabular cup position using x-ray/ct registration,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2000, pp. 1104–1113 (pp. 21, 25).

[82] D. A. LaRose, 'Iterative x-ray/ct registration using accelerated volume rendering,' Ph.D. dissertation, Citeseer, 2001 (pp. 25, 29).

[83] L. Lemieux, R. Jagoe, D. Fish, N. Kitchen and D. Thomas, 'A patient-to-computed-tomography image registration method based on digitally

reconstructed radiographs,' *Medical physics*, vol. 21, no. 11, pp. 1749–1760, 1994 (p. 25).

[84] R. Liao, C. Guetter, C. Xu, Y. Sun, A. Khamene and F. Sauer, 'Learning-based 2D/3D rigid registration using jensen-shannon divergence for image-guided surgery,' in *International Workshop on Medical Imaging and Virtual Reality*, Springer, 2006, pp. 228–235 (pp. 21, 25).

[85] D. Litzenberg, L. A. Dawson, H. Sandler *et al.*, 'Daily prostate targeting using implanted radiopaque markers,' *International Journal of Radiation Oncology* Biology* Physics*, vol. 52, no. 3, pp. 699–703, 2002 (pp. 12, 13).

[86] H. Livyatan, Z. Yaniv and L. Joskowicz, 'Gradient-based 2-d/3-d rigid registration of fluoroscopic x-ray to ct,' *IEEE Transactions on medical imaging*, vol. 22, no. 11, pp. 1395–1406, 2003 (p. 17).

[87] J. Louchet, 'Using an individual evolution strategy for stereovision,' *Genetic Programming and Evolvable Machines*, vol. 2, no. 2, pp. 101–109, Jun. 2001, ISSN: 1573-7632. DOI: 10.1023/A:1011544128842 (p. 36).

[88] J. Lv, M. Yang, J. Zhang and X. Wang, 'Respiratory motion correction for free-breathing 3d abdominal mri using cnn-based image registration: A feasibility study,' *The British journal of radiology*, vol. 91, no. xxxx, p. 20 170 788, 2018 (pp. 11, 23).

[89] S. F. Al-Maliki and F. P. Vidal, *Examples of stochastic optimisation algorithms*, [Online; accessed 08/08/2019], 2019. [Online]. Available: https://github.com/Shatha1978/Optimisation-algorithm-examples (p. 49).

[90] B. Marami, S. S. M. Salehi, O. Afacan *et al.*, 'Temporal slice registration and robust diffusion-tensor reconstruction for improved fetal brain structural connectivity analysis,' *NeuroImage*, vol. 156, pp. 475–488, 2017 (p. 11).

[91] P. Markelj, D. Tomazevic, F. Pernus and B. Likar, 'Robust gradient-based 3-d/2-d registration of ct and mr to x-ray images,' *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1704–1714, 2008 (p. 17).

[92] P. Markelj, D. Tomaževič, B. Likar and F. Pernuš, 'A review of 3d/2d registration methods for image-guided interventions,' *Medical image analysis*, vol. 16, no. 3, pp. 642–661, 2012 (pp. 10, 14).

[93] M. R. Mayberg, E. LaPresto and E. J. Cunningham, 'Image-guided endoscopy: Description of technique and potential applications,' *Neurosurgical Focus*, vol. 19, no. 1, pp. 1–5, 2005 (p. 11).

[94] R. A. McLaughlin, J. Hipwell, D. J. Hawkes, J. A. Noble, J. V. Byrne and T. C. Cox, 'A comparison of a similarity-based and a feature-based 2-D-3-D registration method for neurointerventional use,' *IEEE Transactions on Medical Imaging*, vol. 24, no. 8, pp. 1058–1066, 2005 (p. 25).

[95] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, 'Equation of state calculations by fast computing machines,' *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, Jun. 1953. DOI: 10.1063/1.1699114 (p. 34).

[96] S. Miao, Z. J. Wang, Y. Zheng and R. Liao, 'Real-time 2d/3d registration via cnn regression,' in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2016, pp. 1430–1434 (pp. 21, 22).

[97] S. P. Morozov, A. Andreychenko, N. Pavlov *et al.*, 'Mosmeddata: Chest ct scans with covid-19 related findings dataset,' *arXiv preprint arXiv:2005.06465*, 2020 (p. 111).

[98] S. G. Mueller, M. W. Weiner, L. J. Thal *et al.*, 'The alzheimer's disease neuroimaging initiative,' *Neuroimaging Clinics*, vol. 15, no. 4, pp. 869–877, 2005 (p. 23).

[99] R. Munbodh, Z. Chen, D. A. Jaffray, D. J. Moseley, J. P. Knisely and J. S. Duncan, 'A frequency-based approach to locate common structure for 2D-3D intensity-based registration of setup images in prostate radiotherapy,' *Medical physics*, vol. 34, no. 7, pp. 3005–3017, 2007 (p. 25).

[100] R. Munbodh, D. A. Jaffray, D. J. Moseley *et al.*, 'Automated 2D-3D registration of a radiograph and a cone beam CT using line-segment enhancement,' *Medical physics*, vol. 33, no. 5, pp. 1398–1411, 2006 (pp. 21, 25).

[101] R. Munbodh, H. D. Tagare, Z. Chen *et al.*, '2D-3D registration for prostate radiation therapy based on a statistical model of transmission images,' *Medical physics*, vol. 36, no. 10, pp. 4555–4568, 2009 (pp. 21, 25).

[102] M. J. Murphy, 'An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery,' *Medical Physics*, vol. 24, no. 6, pp. 857–866, 1997 (p. 25).

[103] M. J. Murphy, 'The importance of computed tomography slice thickness in radiographic patient positioning for radiosurgery,' *Medical Physics*, vol. 26, no. 2, pp. 171–175, 1999 (p. 25).

[104] Y. Nakajima, Y. Tamura, Y. Sato *et al.*, 'Preoperative analysis of optimal imaging orientation in fluoroscopy for voxel-based 2-D/3-D registration,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2002, pp. 485–492 (pp. 21, 25).

[105] Y. Nakajima, T. Tashiro, N. Sugano *et al.*, 'Fluoroscopic bone fragment tracking for surgical navigation in femur fracture reduction by incorporating optical tracking of hip joint rotation center,' *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1703–1706, 2007 (p. 25).

[106] E. Neumann, S. Lefèvre, B. Zimmermann, S. Gay and U. Müller-Ladner, 'Rheumatoid arthritis progression mediated by activated synovial fibroblasts,' *Trends in molecular medicine*, vol. 16, no. 10, pp. 458–468, 2010 (p. 46).

[107] T. Okabe, Y. Jin and B. Sendhoff, 'A critical survey of performance indices for multi-objective optimisation,' in *The 2003 Congress on Evolutionary Computation, 2003. CEC'03.*, IEEE, vol. 2, 2003, pp. 878–885 (p. 70).

[108] E. Pelechas, E. Kaltsonoudis, P. V. Voulgari and A. A. Drosos, *Illustrated Handbook of Rheumatic and Musculo-Skeletal Diseases*. Springer, 2018, ISBN: 978-3-030-03664-5 (p. 46).

[109] G. P. Penney, P. G. Batchelor, D. L. Hill, D. J. Hawkes and J. Weese, 'Validation of a two-to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images,' *Medical physics*, vol. 28, no. 6, pp. 1024–1032, 2001 (pp. 21, 25).

[110] G. P. Penney, P. J. Edwards, J. H. Hipwell, M. Slomczykowski, I. Revie and D. J. Hawkes, 'Postoperative calculation of acetabular cup position using 2-D–3-D registration,' *IEEE transactions on biomedical engineering*, vol. 54, no. 7, pp. 1342–1348, 2007 (pp. 21, 25).

[111] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill *et al.*, 'A comparison of similarity measures for use in 2-D-3-D medical image registration,' *IEEE transactions on medical imaging*, vol. 17, no. 4, pp. 586–595, 1998 (pp. 21, 25, 27).

[112] T. Peters and K. Cleary, *Image-guided interventions: technology and applications*. Springer Science & Business Media, 2008 (p. 11).

[113] T. M. Peters, 'Image-guidance for surgical procedures,' *Physics in Medicine & Biology*, vol. 51, no. 14, R505, 2006 (p. 11).

[114] K. Pirpinia, P. A. Bosman, C. E. Loo *et al.*, 'The feasibility of manual parameter tuning for deformable breast mr image registration from a multi-objective optimization perspective,' *Physics in Medicine & Biology*, vol. 62, no. 14, p. 5723, 2017 (p. 69).

[115] A. Pommert, K. H. Höhne, B. Pflesser *et al.*, 'Creating a high-resolution spatial/symbolic model of the inner organs based on the Visible Human,' *Medical Image Analysis*, vol. 5, no. 3, pp. 221–228, 2001, ISSN: 1361-8415. DOI: 10.1016/S1361-8415(01)00044-5 (p. 90).

[116] X. S. Qi, 'Image-guided radiation therapy,' in *An Introduction to Medical Physics*, Springer, 2017, pp. 131–173 (p. 11).

[117] P. Rajpurkar, J. Irvin, A. Bagul *et al.*, 'Mura: Large dataset for abnormality detection in musculoskeletal radiographs,' *arXiv preprint arXiv:1712.06957*, 2017 (pp. 5, 8, 52).

[118] K. S. Rhode, D. L. Hill, P. J. Edwards *et al.*, 'Registration and tracking to integrate x-ray and mr images in an xmr facility,' *IEEE transactions on medical imaging*, vol. 22, no. 11, pp. 1369–1378, 2003 (pp. 20, 21).

[119] M.-M. Rohé, M. Datar, T. Heimann, M. Sermesant and X. Pennec, 'Svf-net: Learning deformable image registration using shape matching,' in *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*, Springer, 2017, pp. 266–274 (p. 23).

[120] T. Rohlfing, J. Denzler, C. Grassl, D. B. Russakoff and C. R. Maurer, 'Markerless real-time 3-D target region tracking by motion backprojection from projection images,' *IEEE transactions on medical imaging*, vol. 24, no. 11, pp. 1455–1468, 2005 (pp. 21, 25).

[121] T. Rohlfing, D. B. Russakoff, J. Denzler, K. Mori and C. R. Maurer Jr, 'Progressive attenuation fields: Fast 2D-3D image registration without precomputation,' *Medical Physics*, vol. 32, no. 9, pp. 2870–2880, 2005 (pp. 21, 25).

[122] P. Romanelli and A. Conti, 'Image-guided stereotactic radiosurgery for spinal pathology,' in *Tumors of the Central Nervous System, Volume 12*, Springer, 2014, pp. 321–333 (p. 11).

[123] P. Romanelli, D. W. Schaal and J. R. Adler, 'Image-guided radiosurgical ablation of intra-and extra-cranial lesions,' *Technology in cancer research & treatment*, vol. 5, no. 4, pp. 421–428, 2006 (p. 11).

[124] D. B. Russakoff, T. Rohlfing, J. R. Adler Jr and C. R. Maurer Jr, 'Intensity-based 2D-3D spine image registration incorporating a single fiducial marker1,' *Academic radiology*, vol. 12, no. 1, pp. 37–50, 2005 (pp. 21, 25).

[125] D. B. Russakoff, T. Rohlfing, A. Ho *et al.*, 'Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data,' in *International Workshop on Biomedical Image Registration*, Springer, 2003, pp. 151–160 (pp. 21, 25).

[126] D. B. Russakoff, T. Rohlfing, K. Mori *et al.*, 'Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration,' *IEEE transactions on medical imaging*, vol. 24, no. 11, pp. 1441–1454, 2005 (pp. 21, 25).

[127] O. Sadowsky, G. Chintalapani and R. H. Taylor, 'Deformable 2D-3D registration of the pelvis with a limited field of view, using shape statistics,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2007, pp. 519–526 (pp. 20, 21, 68).

[128] O. Sadowsky, K. Ramamurthi, L. M. Ellingsen, G. Chintalapani, J. L. Prince and R. H. Taylor, 'Atlas-assisted tomography: Registration of a deformable atlas to compensate for limited-angle cone-beam trajectory,' in *3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.*, IEEE, 2006, pp. 1244–1247 (pp. 20, 21, 68).

[129] S. S. M. Salehi, S. Khan, D. Erdogmus and A. Gholipour, 'Real-time deep pose estimation with geodesic loss for image-to-template rigid registration,' *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 470–481, 2018 (p. 22).

[130] D. Sarrut and S. Clippe, 'Geometrical transformation approximation for 2D/3D intensity-based registration of portal images and CT scan,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2001, pp. 532–540 (p. 25).

[131] G. Schett, 'Erosive arthritis,' *Arthritis research & therapy*, vol. 9, no. 1, pp. 1–6, 2007 (pp. 46, 47).

[132] J. Schindelin, I. Arganda-Carreras, E. Frise *et al.*, 'Fiji: An open-source platform for biological-image analysis,' *Nature methods*, vol. 9, no. 7, pp. 676–682, 2012. DOI: 10.1038/nmeth.2019 (p. 50).

[133] W. Schneider, T. Bortfeld and W. Schlegel, 'Correlation between ct numbers and tissue parameters needed for monte carlo simulations of clinical dose distributions,' *Physics in Medicine & Biology*, vol. 45, no. 2, p. 459, 2000 (p. 43).

[134] F. Schulze, K. Bühler, A. Neubauer, A. Kanitsar, L. Holton and S. Wolfsberger, 'Intra-operative virtual endoscopy for image guided endonasal transsphenoidal pituitary surgery,' *International journal of computer assisted radiology and surgery*, vol. 5, no. 2, pp. 143–154, 2010 (p. 11).

[135] J. Senthilnath, S. Omkar, V. Mani and T. Karthikeyan, 'Multi-objective optimization of satellite image registration using discrete particle swarm optimisation,' in *2011 Annual IEEE India Conference*, IEEE, 2011, pp. 1–5 (p. 69).

[136] Y. Seppenwoolde, H. Shirato, K. Kitamura *et al.*, 'Precise and real-time measurement of 3d tumor motion in lung due to breathing and heartbeat, measured during radiotherapy,' *International Journal of Radiation Oncology* Biology* Physics*, vol. 53, no. 4, pp. 822–834, 2002 (p. 12).

[137] D. W. Shattuck, M. Mirza, V. Adisetiyo *et al.*, 'Construction of a 3d probabilistic atlas of human cortical structures,' *Neuroimage*, vol. 39, no. 3, pp. 1064–1080, 2008 (p. 23).

[138] H. Shirato, S. Shimizu, K. Kitamura *et al.*, 'Four-dimensional treatment planning and fluoroscopic real-time tumor tracking radiotherapy for moving tumor,' *International*

Journal of Radiation Oncology Biology* Physics*, vol. 48, no. 2, pp. 435–442, 2000 (p. 12).

[139] M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab and N. Komodakis, 'A deep metric for multimodal registration,' in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part III 19*, Springer, 2016, pp. 10–18 (p. 21).

[140] L. Sirois, D. Hristov and B. Fallone, 'Three-dimensional anatomy setup verification by correlation of orthogonal portal images and digitally reconstructed radiographs,' *Medical physics*, vol. 26, no. 11, pp. 2422–2428, 1999 (p. 25).

[141] J. M. Sloan, K. A. Goatman and J. P. Siebert, 'Learning rigid image registration-utilizing convolutional neural networks for medical image registration,' 2018 (p. 22).

[142] H. Sokooti, B. De Vos, F. Berendsen, B. P. Lelieveldt, I. Išgum and M. Staring, 'Nonrigid image registration using multi-scale 3d convolutional neural networks,' in *Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20*, Springer, 2017, pp. 232–239 (p. 23).

[143] C. Studholme, D. L. Hill and D. J. Hawkes, 'An overlap invariant entropy measure of 3d medical image alignment,' *Pattern recognition*, vol. 32, no. 1, pp. 71–86, 1999 (p. 26).

[144] A. Sujar, A. Meuleman, P.-F. Villard, M. García and F. P. Vidal, 'gVirtualXRay: Virtual X-Ray Imaging Library on GPU,' in *Computer Graphics and Visual Computing (CGVC)*, The Eurographics Association, 2017, ISBN: 978-3-03868-050-5. DOI: 10.2312/cgvc.20171279 (p. 43).

[145] T. S. Tang, R. E. Ellis and G. Fichtinger, 'Fiducial registration from a single x-ray image: A new technique for fluoroscopic guidance and radiotherapy,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2000, pp. 502–511 (pp. 12, 13).

[146] T. S. Tang, N. MacIntyre, H. Gill *et al.*, 'Accurate assessment of patellar tracking using fiducial and intensity-based fluoroscopic techniques,' *Medical image analysis*, vol. 8, no. 3, pp. 343–351, 2004 (p. 25).

[147] D. Tomazevic, B. Likar and F. Pernus, '3-D/2-D registration by integrating 2-D information in 3-D,' *IEEE transactions on medical imaging*, vol. 25, no. 1, pp. 17–27, 2005 (p. 20).

[148] D. Tomazevic, B. Likar, T. Slivnik and F. Pernus, '3-d/2-d registration of ct and mr to x-ray images,' *IEEE transactions on medical imaging*, vol. 22, no. 11, pp. 1407–1416, 2003 (p. 17).

[149] G. J. Tornai, G. Cserey and I. Pappas, 'Fast drr generation for 2d to 3d registration on gpus,' *Medical physics*, vol. 39, no. 8, pp. 4795–4799, 2012 (p. 111).

[150] M. Q. Tran, T. Do, H. Tran, E. Tjiputra, Q. D. Tran and A. Nguyen, 'Light-weight deformable registration using adversarial learning with distilling knowledge,' *IEEE transactions on medical imaging*, vol. 41, no. 6, pp. 1443–1453, 2022 (p. 23).

[151] A. Valsecchi, E. Bermejo, S. Damas and O. Cordón, 'Metaheuristics for medical image registration,' in *Handbook of Heuristics*, R. Martí, P. Panos and M. G. C. Resende, Eds., Cham: Springer International Publishing, 2017, pp. 1–22, ISBN: 978-3-319-07153-4 (pp. 38, 53).

[152] D. A. Van Veldhuizen, *Multiobjective evolutionary algorithms: classifications, analyses, and new innovations*. Air Force Institute of Technology, 1999 (p. 70).

[153] T. Vercauteren, X. Pennec, A. Perchant and N. Ayache, 'Diffeomorphic demons: Efficient non-parametric image registration,' *NeuroImage*, vol. 45, no. 1, S61–S72, 2009 (p. 23).

[154] M. Vermandel, N. Betrouni, J.-Y. Gauvrit, D. Pasquier, C. Vasseur and J. Rousseau, 'Intrinsic 2D/3D registration based on a hybrid approach: Use in the radiosurgical imaging process,' *Cellular and Molecular Biology*, vol. 52, no. 6, pp. 44–53, 2006 (pp. 20, 21, 68).

[155] S. Vetter, I. Mühlhäuser, J. von Recum, P.-A. Grützner and J. Franke, 'Validation of a virtual implant planning system (vips) in distal radius fractures,' in *Orthopaedic Proceedings*, Bone & Joint, vol. 96, 2014, pp. 50–50 (p. 22).

[156] F. P. Vidal, *gVirtualXRay: Virtual X-Ray Imaging Library on GPU*, [Online; accessed 08/08/2019], 2013. [Online]. Available: `https://sourceforge.net/projects/gvirtualxray/` (p. 43).

[157] F. P. Vidal, M. Garnier, N. Freud, J. M. Létang and N. W. John, 'Simulation of X-ray attenuation on the GPU,' in *Proceedings of Theory and Practice of Computer Graphics 2009*, Cardiff, UK: Eurographics Association, Jun. 2009, pp. 25–32, ISBN: 978-3-905673-71-5. DOI: 10.2312/LocalChapterEvents/TPCG/TPCG09/025-032 (pp. 8, 43).

[158] F. P. Vidal, M. Garnier, N. Freud, J. M. Létang and N. W. John, 'Accelerated deterministic simulation of X-ray attenuation using graphics hardware,' in *Eurographics 2010 - Posters*, Eurographics Association, May 2010, Poster 5011. DOI: 10.2312/egp.20101026 (p. 43).

[159] F. P. Vidal, I. T. Mitchell and J. M. Létang, 'Use of fast realistic simulations on gpu to extract cad models from microtomographic data in the presence of strong ct artefacts,' *Precision Engineering*, vol. 74, pp. 110–125, 2022 (p. 7).

[160] F. P. Vidal and P.-F. Villard, 'Simulated motion artefact in computed tomography,' in *Eurographics Workshop on Visual Computing for Biology and Medicine*, The Eurographics Association, 2015, pp. 213–214, ISBN: 978-3-905674-82-8. DOI: 10.2312/vcbm.20151228 (p. 43).

[161] F. P. Vidal and P.-F. Villard, 'Development and validation of real-time simulation of X-ray imaging with respiratory motion,' *Computerized Medical Imaging and Graphics*, vol. 49, pp. 1–15, Apr. 2016. DOI: 10.1016/j.compmedimag.2015.12.002 (pp. 8, 43).

[162] D. Wang, H. Li, X. Wei and X.-P. Wang, 'An efficient iterative thresholding method for image segmentation,' *Journal of Computational Physics*, vol. 350, pp. 657–667, 2017 (p. 111).

[163] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, 'Image quality assessment: From error visibility to structural similarity,' *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004. DOI: 10.1109/TIP.2003.819861 (p. 91).

[164] J. Weese, R. Goecke, G. P. Penney, P. Desmedt, T. M. Buzug and H. Schumann, 'Fast voxel-based 2D/3D registration algorithm using a volume rendering method based on the shear-warp factorization,' in *Medical Imaging 1999: Image Processing*, International Society for Optics and Photonics, vol. 3661, 1999, pp. 802–810 (pp. 21, 25).

[165] J. Weese, T. M. Buzug, C. Lorenz and C. Fassnacht, 'An approach to 2D/3D registration of a vertebra in 2D x-ray fluoroscopies with 3D CT images,' in *CVRMed-MRCAS'97*, Springer, 1997, pp. 119–128 (pp. 21, 25, 27).

[166] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. Hill and D. J. Hawkes, 'Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery,' *IEEE transactions on information technology in biomedicine*, vol. 1, no. 4, pp. 284–293, 1997 (pp. 21, 25).

[167] W. Wein, B. Roeper and N. Navab, '2d/3d registration based on volume gradients,' in *Medical Imaging 2005: Image Processing*, SPIE, vol. 5747, 2005, pp. 144–150 (p. 17).

[168] U. Wickramasinghe, E. Remelli, G. Knott and P. Fua, 'Voxel2mesh: 3d mesh model generation from volumetric data,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2020, pp. 299–308 (p. 111).

[169] Wikipedia contributors, *Cma-es*, [Online; accessed 22-August-2022], 2022. [Online]. Available: `https://en.wikipedia.org/wiki/CMA-ES` (p. 39).

[170] C. J. Willmott and K. Matsuura, 'Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance,' *Climate research*, vol. 30, no. 1, pp. 79–82, 2005 (p. 31).

[171] G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao and D. Shen, 'Unsupervised deep feature learning for deformable registration of mr brain images,' in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part II 16*, Springer, 2013, pp. 649–656 (p. 21).

[172] J. Wu, E. E. A. Fatah and M. R. Mahfouz, 'Fully automatic initialization of two-dimensional–three-dimensional medical image registration using hybrid classifier,' *Journal of Medical Imaging*, vol. 2, no. 2, pp. 024 007–024 007, 2015 (p. 11).

[173] P. Wunsch and G. Hirzinger, 'Registration of cad-models to images by iterative inverse perspective matching,' in *Proceedings of 13th International Conference on Pattern Recognition*, IEEE, vol. 1, 1996, pp. 78–83 (p. 15).

[174] E. Xevi and S. Khan, 'A multi-objective optimisation approach to water management,' *Journal of environmental management*, vol. 77, no. 4, pp. 269–277, 2005 (p. 69).

[175] J. Yao and R. Taylor, 'Assessing accuracy factors in deformable 2D/3D medical image registration using a statistical pelvis model,' in *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, ser. ICCV '03, USA: IEEE Computer Society, 2003, pp. 1329–1334, ISBN: 0769519504 (pp. 4, 19, 21, 68).

[176] B. Yetkin Ekren, 'A multi-objective optimisation study for the design of an avs/rs warehouse,' *International Journal of Production Research*, vol. 59, no. 4, pp. 1107–1126, 2021 (p. 69).

[177] T. S. Yoo, *Insight into images: principles and practice for segmentation, registration, and image analysis*. AK Peters Ltd, 2004, pp. 237–338 (p. 10).

[178] B.-M. You, P. Siy, W. Anderst and S. Tashman, 'In vivo measurement of 3-D skeletal kinematics from sequences of biplane radiographs: Application to knee kinematics,' *IEEE transactions on medical imaging*, vol. 20, no. 6, pp. 514–525, 2001 (pp. 21, 25).

[179] Q. Zhao, C.-R. Chou, G. Mageras and S. Pizer, 'Local metric learning in 2d/3d deformable registration with application in the abdomen,' *IEEE transactions on medical imaging*, vol. 33, no. 8, pp. 1592–1600, 2014 (p. 11).

[180] G. Zheng, 'Effective incorporation of spatial information in a mutual information based 3D-2D registration of a CT volume to x-ray images,' in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2008, pp. 922–929 (pp. 21, 25).

[181] G. Zheng, M. Á. Ballester, M. Styner and L.-P. Nolte, 'Reconstruction of patient-specific 3d bone surface from 2d calibrated fluoroscopic images and point distribution model,' in *International conference on medical image computing and computer-assisted intervention*, Springer, 2006, pp. 25–32 (p. 17).

[182] G. Zheng, X. Zhang, S. Jonić, P. Thévenaz, M. Unser and L.-P. Nolte, 'Point similarity measures based on mrf modeling of difference images for spline-based 2D-3D rigid registration of x-ray fluoroscopy to CT images,' in *International Workshop on Biomedical Image Registration*, Springer, 2006, pp. 186–194 (pp. 21, 25).

[183] L. Zollei, E. Grimson, A. Norbash and W. Wells, '2D-3D rigid registration of x-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators,' in *Proceedings of the 2001 IEEE Computer Society Conference on Computer*

Vision and Pattern Recognition. CVPR 2001, IEEE, vol. 2(11), 2001, pp. 696–703 (pp. 21, 25).