

REGISTRATION OF MEDICAL
IMAGES FOR APPLICATIONS IN
MINIMALLY INVASIVE
PROCEDURES

BOGDAN M. MARIS

DOTTORATO DI RICERCA IN INFORMATICA - XXVI CICLO

PHD PROGRAM IN COMPUTER SCIENCE

2014

University of Verona

Supervisor: prof. Paolo Fiorini

Contents

Abstract	13
Dedication	14
Acknowledgements	15
1 Introduction	16
1.1 What is image registration?	16
1.2 Image registration in the OR: challenges and applications in minimally invasive procedures	17
1.3 Contribution of this work	18
1.4 Thesis outline	18
2 Medical image registration framework	20
2.1 Introduction in medical image registration	20
2.2 Medical image registration components and methodology	23
2.2.1 Distance measures	24
2.2.2 Parametric and non parametric solutions	28
2.2.3 Optimization	30
2.3 Experimental set-up for testing	31
2.4 Why the medical image registration framework cannot work for OR applications?	35
3 Registration of binary images	37
3.1 Introduction	37
3.1.1 Extraction of landmarks and binary images	38
3.1.2 Correspondences and transformation	39
3.2 Test on phantom data	41

3.3	Discussion	43
3.3.1	Why the binary (landmark) registration works better than the intensity-based registration for an OR application?	43
3.3.2	What are the limitations of the binary registration for an OR application?	44
3.3.3	What is the direction in which we should proceed?	44
3.4	A real-time 3D set-up in the OR	44
3.4.1	Realistic phantom set-up	46
3.4.2	Real-time US 3D acquisition and image segmentation	48
3.4.3	Real-time surface registration	50
3.4.4	Experimental test and results	52
3.5	Comments	53
4	Generalized Shape Spaces	55
4.1	Correspondences, registration and the Procrustes method	56
4.2	Shape-spaces and shape-manifolds: intro	60
4.3	Theoretical foundations	61
4.4	Distance distribution and generalized shapes: existence	67
4.5	Distance distribution and generalized shapes: isometric transformation	70
4.6	Distance distribution and generalized shapes: correspondences and registration	77
4.6.1	Distance distribution permutation	77
4.6.2	Determining correspondences	80
4.6.3	Correspondence test and results	84
4.6.4	Registration tests and results	87
4.6.5	Correspondences and registration evaluation using pulmonary landmark points derived from 4D CT image data	99
4.6.6	Conclusions and future work	101
5	A needle guidance system for percutaneous procedures	104
5.1	Introduction	105
5.2	Background	106
5.3	System Work ow	108
5.3.1	Hardware Setup	109

5.3.2	Needle Calibration	110
5.3.3	Needle Navigation	112
5.4	Validation and Results	114
5.4.1	Calibration Accuracy and Precision	114
5.4.2	In-vitro evaluation of the navigation system	115
5.5	Conclusions and Future Work	118
6	Final words and future works	120
	Bibliography	123
A	Image registration results	136
A.1	Parametric registration using mutual information	136
A.2	Parametric registration using normalized gradient field	139
A.3	Multilevel parametric and non parametric registration	141
A.3.1	Input	141
A.3.2	Parametric multilevel spline registration	142
A.3.3	Non parametric multilevel registration	143
A.4	Computation time	148
B	Landmark-based registration results	149
B.1	Landmark-based registration: linear	149
B.2	Landmark-based registration: non-linear	152
B.2.1	Quadratic registration	152
B.2.2	Thin-plate-spline registration	154

List of Tables

3.1	Accuracy and the precision of the system, both for the GRE and TRE, as the average and the standard deviation of the error. The last column gives the average time taken by the entire system to accomplish the registration.	53
4.1	Correspondences test	86
4.2	Comparison of Error Values after the registration using our algorithm and the other two approaches.	93
4.3	Percentage of good correspondences for each patient and different breathing phases.	100
4.4	The Hausdorff distance after the rigid and the non rigid registration. . .	101
5.1	Results for the needle position and orientation calibration.	114
5.2	Comparison of Error Values in the two experimental setups.	117
5.3	Comparison of Timing in the two experimental setups.	117
A.1	Computation time of the principal registration algorithm tested.	148

List of Figures

2.1	A counterclockwise rotation will move clockwise the template image when the transformation is done using the (2.3) equation	24
2.2	CIRS 057 interventional 3D abdominal phantom.	31
2.3	Different scannings of the phantom. From left to right: MRI, CT, US.	32
2.4	The images used in the registration process. On the left hand side the US template image; on the right hand side the CT reference (x e d) image.	32
2.5	Internal anatomy of the anatomical phantom.	33
2.6	From left to right: the 3D image of the phantom with 4 markers that identify its position in space; a CT slice of the phantom; the same CT slice with the outline of the phantom in transparency.	34
2.7	(a) The optical tracking system de nes the global refere nce system. (b) The US probe and the transformations involved in the calibration process. (c) The calibration of the US probe is based on the identi cation of a xed point in different poses of the probe.	34
3.1	The template and the reference images with corresponding landmarks.	38
3.2	The CAD model for the realization of the phantom (left) and the pancreas phantom (right) with the internal duct.	46
3.3	Top: A US slice before (left) and after the segmentation process (right), where the blue line is the organ surface, the green line is cyst's edge and the red line is edge of a duct. Bottom: 3D reconstruction of the segmented surface (blue) with duct (red) and cyst (green).	48
3.4	Real time US image segmentation pipeline.	49
3.5	Left: CAD extracted surface. Right: US segmented image (blue) registered with the CAD surface (white).	50
3.6	Block diagram of the proposed framework	52

4.1	The point sets $\{A, B, C, D\}$, $\{A, B, C, E\}$, and $\{A, B, C, F\}$ have the same distribution of distances.	66
4.2	Testing 2D data for empirical robustness evaluation. On each of the images the points marked with blue '*' represent the model (joined by a continuous line obtained through interpolation) meanwhile the points marked with red 'o' represent the model transformed by a random permutation and different levels of noise. The first model on the left column follows [53], [82].	84
4.3	Two 3D datasets with different noise levels. The points marked with blue '*' represent the model and the points marked with red 'o' represents the model transformed by a random permutation and different levels of noise.	85
4.4	The Hausdorff distance	89
4.5	The datasets used in the comparison between different algorithms for the registration. (a) The shape rotated by 20 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%). (b) The shape rotated by 40 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%). (c) The shape rotated by 60 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%). (d) The shape rotated by 80 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%)	90
4.6	The registration using our registration algorithm. On each of the figure the blue stars are the fixed points, meanwhile the red circles represent the moving point-set after the registration took place. The Hausdorff distance is represented by the line between the black squares and its value is given in the table 4.2	94

- 4.7 The algorithm [82] applied to the sh shape dataset at noise level 1. (a), (b), (c), (e) represent the registration using the non rigid algorithm, (d), (f) the results using the rigid algorithm. Top row: the Hausdorff distance between the two shapes after the registration (black line between the black squares), also reported in the table 4.2. Bottom row: the fixed dataset is given by the blue stars, meanwhile the red circles represent the moving dataset. The lines give the correspondences between the fixed and the moving points. Observe that the correspondences in (c), (e), (f) are wrong (bottom row) and so is the registration in these cases (top row). 95
- 4.8 The algorithm [82] applied to the sh shape dataset at noise level 3. (a), (b), (c), (e) represent the registration using the non rigid algorithm, (d), (f) the results using the rigid algorithm. Top row: the Hausdorff distance between the two shapes after the registration (black line between the black squares), also reported in the table 4.2. Bottom row: the fixed dataset is given by the blue stars, meanwhile the red circles represent the moving dataset. The lines give the correspondences between the fixed and the moving points. Observe that the correspondences in (c), (e), (f) are wrong (bottom row) and so is the registration in these cases (top row). 96
- 4.9 The algorithm [82] applied to the sh shape dataset at noise level 5. (a), (b), (c), (e) represent the registration using the non rigid algorithm, (d), (f) the results using the rigid algorithm. Top row: the Hausdorff distance between the two shapes after the registration (black line between the black squares), also reported in the table 4.2. Bottom row: the fixed dataset is given by the blue stars, meanwhile the red circles represent the moving dataset. The lines give the correspondences between the fixed and the moving points. Observe that the correspondences in (c), (e), (f) are wrong (bottom row) and so is the registration in these cases (top row). 97
- 4.10 Tests applying the algorithm [53]. The fixed points are represented by blue stars, the moving points by thin red circles before the registration and by bold red circles after the non rigid registration. The lines give the movement of the correspondent red circles. The numerical values of the Hausdorff distance are given in the table 4.2. 98

4.11	A 3D rendering of thoracic image with the pulmonary area and the extracted landmarks overlying a grayscale CT slice.	99
4.12	The 2D grayscale CT slice with 3D landmarks.	100
5.1	A standard 18 gauge biopsy needle equipped with the compact navigation system.	106
5.2	The needle navigation scheme. The plane π is perpendicular on the needle axis. The vector $\mathbf{p} \in \pi$ is the direction toward the entry point. The vector $\mathbf{o} \in \pi$ is the direction toward the correct orientation. . . .	108
5.3	State machine diagram. The threshold error depends on the state: $\epsilon_1 < 1mm, \epsilon_2 < 3mm, \epsilon_3 < 4mm$. The value of the ϵ_θ depends on the accuracy of the orientation we want to reach. The value of ϵ_3 depends on the accuracy we want to reach and should consider the calibration and the tracking errors.	109
5.4	The coordinate systems rigid transformations. The matrix $\begin{matrix} Marker \\ Needle \end{matrix} T$ is the calibration matrix. The origin of the needle coordinate system is translated in the needle tip.	110
5.5	Cross calibration system.	111
5.6	State machine display a) positioning state. b)orientation state. c) insertion step. d) calibration step.	112
5.7	Experimental setup no. 2 with both the PC monitor and the display attached to the needle.	116
A.1	Af ne registration using MI. Top row, from left to right: the reference image, the target image before the rst iteration, the target t image after 44 iterations. Bottom row, from left to right: the target image after the registration with the deformation grid superimposed, target and reference difference image before registration, target and reference difference image after registration.	136
A.2	Af ne registration using MI. The graphical trend of the o bjective function J during 44 iterations (the black line). The green line gives the graph of $ \nabla J $, while the red shows line the graph of the difference between the previous J value and the current value.	137

A.3	Spline registration using MI. Top row, from left to right: the reference image, the target image before the first iteration, the target image after 100 iterations. Bottom row, from left to right: the target image after the registration with the deformation grid superimposed, target and reference difference image before registration, target and reference difference image after registration.	137
A.4	Spline registration using MI. The graphical trend of the objective function J during 100 iterations (the black line). The green line gives the graph of $ \nabla J $, while the red shows line the graph of the difference between the previous J value and the current value.	138
A.5	Affine registration using NGF. The results are obtained after 228 iterations. The scheme is the same as in figure A.2	139
A.6	Affine registration using NGF. The graphical trend of the objective function J during 228 iterations (the black line). The green line gives the graph of $ \nabla J $, while the red line shows the graph of the difference between the previous J value and the current value.	139
A.7	Spline registration using NGF. Top row, from left to right: the reference image, the target image before the first iteration, the target image after 100 iterations. Bottom row, from left to right: the target image after the registration with the deformation grid superimposed, target and reference difference image before registration, target and reference difference image after registration.	140
A.8	Spline registration using NGF. The graphical trend of the objective function J during 100 iterations (the black line). The green line gives the graph of $ \nabla J $, while the red shows line the graph of the difference between the previous J value and the current value.	140
A.9	The target (top row) and the reference (bottom row) images represented using 6 different levels.	141
A.10	The initialization of the algorithm at the first level considered uses the affine transformation.	143
A.11	First level of registration using the affine transformation as initialization and elastic registration to refine.	143
A.12	Second level of registration: the initialization uses the previous obtained transformation.	144

A.13	Third level of registration: the initialization uses the previous obtained transformation.	144
A.14	Fourth level of registration: the initialization uses the previous obtained transformation.	145
A.15	Last level of registration: the initialization uses the previous obtained transformation.	145
A.16	Iteration history of multilevel non-parametric elastic registration: vertical lines separate different levels. The squares represent the initial value of MI, while the crosses show the value on each iteration.	146
A.17	The result of the registration using the curvature.	147
A.18	Iteration history of multilevel non-parametric curvature registration: vertical lines separate different levels. The squares represents the initial value of MI, while the crosses show the value on each iteration.	147
B.1	Overlapped images before registration: target image (US) is represented by the red layer, source image (CT) by the green layer.	149
B.2	The template and the reference images with 3 corresponding pair of landmarks.	150
B.3	The template image before and after the linear registration with the registered landmarks highlighted.	150
B.4	From left to right: a close-up of the linearly registered landmarks, the target (red) and reference (green) images superimposed after the linear registration and a close-up of the same image.	151
B.5	The template and the reference images with 12 corresponding landmarks.	151
B.6	The template image before and after the registration with the registered landmarks highlighted.	152
B.7	From left to right: a close-up of the registered landmarks, the target and reference images superimposed after the registration, a close-up of the superimposed images.	152
B.8	The target and reference image with 6 selected landmarks. On the right the target image registered and the 6 landmarks perfectly aligned by the quadratic model.	153
B.9	The two images overlapped after the quadratic registration with 6 landmarks in the middle and a close-up on the right. On the left a close-up of some registered landmarks.	153

B.10	The two images overlapped after the quadratic registration with 15 landmarks in the middle and a close-up on the right. On the left a close-up of some registered landmarks.	154
B.11	The target and reference image with 15 selected landmarks. On the right the target image registered and the 15 landmarks aligned by the quadratic model.	154
B.12	The target (left) and reference (right) image with 6 selected landmarks.	155
B.13	The target image from g.B.12 registered for $\alpha = 0$	155
B.14	The two images from g.B.12 overlapped after the TPS registration for $\alpha = 0$	156
B.15	The target image from g.B.12 registered for $\alpha = 10$	156
B.16	The two images from g.B.12 overlapped after the TPS registration for $\alpha = 10$	157
B.17	The target image from g.B.12 registered for $\alpha = 100$	157
B.18	The two images from g.B.12 overlapped after the TPS registration for $\alpha = 100$	158
B.19	The target image from g.B.12 registered for $\alpha = 1000$	158
B.20	The two images from g.B.12 overlapped after the TPS registration for $\alpha = 1000$	159
B.21	The target (left) and reference (right) image with 15 selected landmarks.	159
B.22	The target image from g.B.21 registered for $\alpha = 0$	160
B.23	The two images from g.B.21 overlapped after the TPS registration for $\alpha = 0$	160
B.24	The target image from g.B.21 registered for $\alpha = 10$	161
B.25	The two images from g.B.21 overlapped after the TPS registration for $\alpha = 100$	161
B.26	The target image from g.B.21 registered for $\alpha = 100$	162
B.27	The two images from g.B.21 overlapped after the TPS registration for $\alpha = 100$	162
B.28	The target image from g.B.21 registered for $\alpha = 1000$	163
B.29	The two images from g.B.21 overlapped after the TPS registration for $\alpha = 1000$	163

Abstract

REGISTRATION OF MEDICAL IMAGES FOR APPLICATIONS IN MINIMALLY INVASIVE PROCEDURES

The registration of medical images is necessary to establish spatial correspondences across two or more images. Registration is rarely the end-goal, but instead, the results of image registration are used in other tasks.

The starting point of this thesis is to analyze which methods at the state of the art of image registration are suitable to be used in assisting a physician during a minimally invasive procedure, such as a percutaneous procedure performed manually or a teleoperated intervention performed by the means of a robot.

The first conclusion is that, even if much previous work has been devoted to develop registration algorithms to be applied in the medical context, most of them are not designed to be used in the operating room scenario (OR) because, compared to other applications, the OR requires also a strong validation, real-time performance and the presence of other instruments. Almost all of these algorithms are based on a three phase iteration: optimize-transform-evaluate similarity. In this thesis, we study the feasibility of this three steps approach in the OR, showing the limits that such approach encounter in the applications we are considering. We investigate how could a simple method be realizable and what are the assumptions for such a method to work. We then develop a theory that is suitable to register large sets of unstructured data extracted from medical images keeping into account the constraints of the OR.

The use of the whole radiologic information is not feasible in the OR context, therefore the method we are introducing registers processed dataset extracted from the original medical images.

The framework we propose is designed to find the spatial correspondence in closed form keeping into account the type of the data, the real-time constraint and the presence of noise and/or small deformations. The theory and algorithms we have developed are in the framework of the shape theory proposed by Kendall in [58] and uses a global descriptor of the shape to compute the correspondences and the distance between shapes.

Since the registration is only a component of a medical application, the last part of the thesis is dedicated to some practical applications in the OR that can benefit from the registration procedure.

Dedication

To my beloved Elena, to Eric and Cristian, the joys of my life.

Acknowledgements

I am most grateful to my advisor Professor Paolo Fiorini for his support and guidance and for having created the Altair robotics research group.

His way of leading the laboratory and his dedication is one of the most important resource that I and my colleagues have benefited.

It was really a pleasure to have such a personality as supervisor during my doctorate.

Encouraged by Professor Fiorini I had the opportunity to participate at a number of international conferences and some very important summer schools that gave me also the opportunity to visit beautiful places all over the world. The years I spent in the laboratory were to me not only years of great intellectual ferment, but also years of major transformation in personal life.

I found excellent interlocutors in the laboratory and the discussions with them proved to be useful not only in the scientific perspective but also in many aspects of every day life.

I am very indebted first of all to Diego who provided me all the data I needed and who was a very good collaborator in working and writing research papers and project deliverables, to Davide who was always available to give me advices from his Ph.D experience and to Otar for its personal support and generosity.

I remember with pleasure the delicious discussions I had with my colleagues during not so delicious lunches in the canteen, especially with Andrea, Luisa, Francesco, Giovanni and Giacomo.

Meeting so many great people in the lab was a privilege for me and I want to remember here most of them: Fabrizio, Michele, Michela, Lorenza, Carlo, Debora, Lorenzo, Marco, Marta, Riccardo, Francesco, Luca, Andrea, ...

Finally, I want to thank my family for the education I received and for their unconditional support in all the choices I made.

Chapter 1

Introduction

Wir müssen wissen, wir werden
wissen.

David Hilbert

1.1 What is image registration?

Image registration is a method to align two or more images. Most of the times images of the same scene are obtained at different times, from different perspectives or using different modalities. Image registration is required to find the correspondences between the imaged objects and to recover the geometrical transformation that aligns the images, so that the aligned images can be directly compared, combined and analyzed. Main applications of image registration include medical imaging, remote sensing and computer vision .

The simplest example of image registration is the rigid alignment where images need to be rotated and shifted with respect to each other to achieve correspondence.

Another class of methods for image registration is the non-rigid transformation, where two images are related through a non-rigid geometric transformation. This need often arises in medical imaging, where an imaged body organ undergoes soft-tissue type deformation. Non-rigid image registration is one of the key technologies in medical image analysis together with image segmentation.

1.2 Image registration in the OR: challenges and applications in minimally invasive procedures

The image registration in the OR is part of a larger system that comprises also other key technologies:

- Medical imaging and low level image processing not comprising registration and segmentation.
- Data visualization.
- Segmentation.
- Tracking systems.
- Medical robots.
- Human Computer Interaction (HCI).

This framework is needed to support minimally invasive techniques by replacing direct visual feedback in the open surgery with indirect, image-based, feedback.

These systems are also designed to mitigate the learning curve for minimally invasive procedures, to reduce the variability of the outcome and to enable the development of new procedures, allowing the physicians to perform procedures that were previously considered too dangerous or not feasible before.

The two fundamental image techniques are the identification of anatomical structures in the images (segmentation) and establishment of the spatial relationship between the imagery and the patient (registration).

The registration is necessary to integrate real anatomical structures with patient medical images for planning and execution, substituting the error prone mental work of the physician.

In order to ensure the specifications of a framework for minimally invasive techniques, we must consider the following constraints in the design of a registration method:

1. accuracy, measured as target registration error (TRE), which indicates how far the predicted position of the anatomical target is from its actual position.
2. speed, or how long does the algorithm takes to produce the solution.

3. robustness, or how well the algorithm deals with noise and outliers.

These constraints depend on the medical task. The accepted accuracy for most of the tasks is of several millimeters, while the speed can range from several seconds to a couple of minutes.

1.3 Contribution of this work

1. A comparative study of the state of the art of registration algorithms in medical image processing as possible solution to the registration problem in the OR: since most of the work in the field of medical image registration is based on the three stages iterative algorithms, we have analyzed the feasibility of this approach in our case study and we have underlined the limits of the current algorithms through a testing process.
2. The development of a theory for point-set registration to be used in the OR applications: we have developed the theoretical basis and implemented algorithms to register dense binary images. This approach is conceived to satisfy the constraints of the OR applications and to substitute the currently used approaches based on iterative algorithms where the correspondences and the transformation are two unrelated stages and may produce arbitrary results.
3. Development of tools to perform minimally invasive procedures in the OR: we have designed a tool to assist the manual insertion of a needle or a probe for diagnostic (biopsy) or therapeutic (ablation) procedures.

1.4 Thesis outline

Chapter 2 presents the classical solution to the registration problem as an optimization of the similarity measure in the parameters space or over the regularization constraint of the physical model adopted. We have tested both the parametric and the non-parametric solutions on images obtained from an anatomical phantom.

In chapter 3, we discuss some of the solutions of the registration currently used in commercial systems for OR interventions and we conduct a phantom study of a simple registration method in order to have the complete control over the registration results both in term of accuracy, robustness and speed of execution.

In chapter 4, we introduce a novel method to solve the point-set registration in closed form. We derive the method, analyze its properties and validate it against other popular methods. We also demonstrate the application to the non-rigid registration of images, that is in the presence of noise or in the case of pulmonary landmark point-sets affected by the breathing motion.

In chapter 5, we present an innovative tool composed of a small screen, a tracking device, a communication board attached to a standard needle or probe to be used in minimally invasive procedures.

Chapter 2

Medical image registration framework

Les mathématiciens n'étudient pas
des objets, mais des relations entre
les objets.

Henri Poincaré

In this chapter we give a brief overview of the state of the art of registration methods in medical image processing and we introduce the classical approaches for the solution of this problem. After the introduction of the registration framework and of the components involved, we study whether these algorithms can be used in the OR or not. The study we have performed employs, as input to the algorithms, multi-modal medical images obtained from an anatomical phantom with some basic structures visible under the ultrasound (US) and computed tomography (CT) scanning.

2.1 Introduction in medical image registration

Image registration is the procedure of aligning two or more images of the same scene taken from different viewpoints, at different time, and/or by different sensors, so that corresponding features can be easily related. Image registration has application to many fields but the one addressed here is medical imaging and medical applications. Lately, medical image registration evolved from an application of medical imaging to a discipline in itself. Image registration has also become one of the more successful areas of image processing, with fully automated algorithms available in a number of applications.

In the case of a surgical procedure, the images to be registered are acquired in two steps: the first one takes place before the procedure and we call such dataset the *pre-operative* image, the second one is during the procedure and is called *intra-operative* dataset. By registering the two dataset, a spatial relationship between the anatomical structures in the two images and with the body of the patient is established and it is also possible to integrate spatial information about physiological functions and pathologies or other abnormalities.

Deriving the correspondence of spatial information in medical images and equivalent structures in the body is fundamental to image interpretation and analysis.

In the classical clinical scenario, the diagnostician's task is to mentally combine the images from several modalities acquired at different times to draw useful clinical conclusions or to plan a surgical intervention. This generally requires mental compensation for changes in subject position. In image guided interventions, image registration establishes correspondence between images and physical space.

Nowadays it is common to obtain images from patients multiple times, either by repeated single modality imaging, or by imaging with different modalities. It is also standard procedure to use sequences of images obtained dynamically from the same patient, often at many frames per second, as in the case of the *ultrasound* (US) image acquisition. This image data acquired make it necessary to relate one image to another in order to assist in extracting relevant clinical information. Image registration can help this task by combining complementary information from different modalities in the multimodal case and by enabling accurate comparisons between images from the same modality.

Important applications of the registration in clinical practice include: the monitoring of changes due to disease progression or treatment, integration of dynamic acquisition for functional studies (e.g. perfusion images) with the anatomical images, image-guided interventions or diagnosis such as tumor ablation or image guided biopsy, in which images acquired before the procedure where the target is identified are integrated/registered with the images acquired during the intervention or with the instruments used during the procedure, enabling the surgeon to guide his or her work.

The images to be registered are in digital form. This means that the medical images are made up of a rectangular array of small elements called *pixels*; each pixel has an associated image intensity value. This array provides the reference system of the image. In the case of *computed tomography* (CT) images a typical slice is composed of 512×512 pixels, and each will correspond to a small planar cut of about 0.5×0.5

mm^2 . The pixel dimension determines the spatial resolution of the image. In the case of CT or *magnetic resonance image* (MRI) the images are stacked together to form a 3D volume, therefore each pixel will correspond to a small volume of tissue, or *voxel*. According to the slice spacing, i.e. the distance between two adjacent slices, the voxel volume is completely determined. For instance, in the case of 1.5 mm spacing and considering the pixel dimension of $0.5 \times 0.5 \text{ mm}^2$, the tissue volume covered by each voxel in the image will be of $0.5 \times 0.5 \times 1.5 \text{ mm}^3$. As in the pixel case, the number stored in each voxel represents the image intensity and corresponds to the integral of some physical attribute measured over this volume.

During an image-guided intervention, the registration is needed to establish the correspondence between the image and the physical space of the patient. This correspondence allows the image to provide a map for the navigation with the goal of making the intervention more accurate, safer and less invasive for the patient. The image registration techniques are already present in the clinical set-up for image-guided neurosurgery systems and in orthopedic surgery.

Considering the nature of the problem in medical image registration, that is the deformability of the tissue encountered in most of the medical images, except the images where the main interest is on rigid structures such as bones, most of the current registration algorithms try to find a solution that involves deformation.

Then, the challenge of this thesis is first to find or design image registration techniques to be included in an image-guided system that involves structures subject to deformations.

In the following we present some of the state of the art solutions and we analyze whether these solutions have or have not the desired characteristics so that they can be employed in an image-guided system.

The stereotactic method, one of the first registration methods used in brain surgery, was developed at the beginning of the 20-th century and is still in use but the modern systems involve high resolution pre-operative 3D dataset integrated with real-time intra-operative images and sometimes also with the surgical tools by the mean of tracking devices.

The pioneer works for the introduction of fully automated algorithms for both multi-modality [127], [43] and inter-modality [128], [119], [26], [109], [110] were presented in the first half of the 1990s. A significant breakthrough for the multi-modal registration was achieved with the introduction of image alignment measures and algorithms based on entropy and, in particular, with the introduction of mutual information

measure derived from the information theory developed by Shannon in 1948 [104].

The modern medical image analysis is focused on algorithms that handle more and more complicated transformations needed to model soft tissue deformation. General reviews of the field may be found in [17], [118], [72], [44], [132], [77], [38], [113], [79], [78], [37], [108].

2.2 Medical image registration components and methodology

Even if there is no general solution, most of the registration algorithms in medical imaging field are based on three components:

1. A distance measure between images to be registered.
2. A transformation model which could be parametric and nonparametric.
3. An optimization method.

In general, registration can be performed on two or more images. Without loss of generality, we may assume that the registration involve only two images. One of the images is referred to as the template image, denoted in the following by the calligraphic character \mathcal{T} , and the other is referred to as the target or reference image, denoted by \mathcal{R} .

Images are considered as mappings from a domain into the real numbers. The domain is denoted by $W \subset \mathbb{R}^d$, where d denotes the spatial dimensionality of the given data.

$$\mathcal{T} : W \rightarrow \mathbb{R}, W \subset \mathbb{R}^d. \quad (2.1)$$

Typically $d=2$ or $d=3$. To each point in the domain, a gray value is assigned.

The goal of the registration is to estimate the transformation f that minimizes an energy of the form

$$\mathcal{D}(\mathcal{T} \circ f, \mathcal{R}) + \mathcal{S}(f), \quad (2.2)$$

where $\mathcal{T} \circ f$ is the transformed template image, \mathcal{D} measures the distance between images or the image similarity, and \mathcal{S} measures the smoothness of the transform aiming also to favor any specific property in the solution that the user requires.

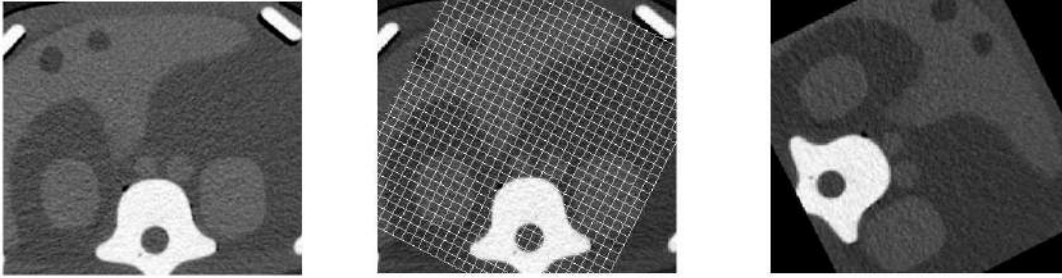


Figure 2.1: A counterclockwise rotation will move clockwise the template image when the transformation is done using the (2.3) equation

The transformation f applied to the template image \mathcal{T} is a vector-valued function $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$ and

$$(\mathcal{T} \circ f)(x) = \mathcal{T}(f(x)) \quad (2.3)$$

This approach, called Eulerian, is important from a practical point of view, since when the locations of the pixels/voxels in the template image are mapped to positions that do not correspond to pixels/voxels in the template image, their intensities can be calculated by interpolating the intensity values of the neighboring pixels/voxels. As a result, the corresponding transformation of the template image is counterintuitive: when the grid is rotated counterclockwise, the image is rotated clockwise (Figure 2.1).

The transformation at every position $x \in W$ may be given in a vector space as the addition of an identity transformation with the displacement field u :

$$f(x) = x + u(x), \quad (2.4)$$

or as a group structure, where the group operation is the function composition and the identity is given by the identity transformation.

2.2.1 Distance measures

The objective function (see equation (2.2)) is the sum of a measure of distance between the transformed template image and the reference image and the smoothing functional S over the mapping f . We focus here on the distance measure, the smoothing functional will be discussed in the next subsection.

In the case where the same anatomical structures are assumed to correspond to similar intensity values, correlation based matching produces dense depth maps by

calculating the disparity at each pixel/voxel within a neighborhood.

The simplest distance measures in this case are given by the L_p -norms of the intensity differences. The most used norms in image registration are the L_1 norm or *sum of absolute differences* SAD (2.5) and the L_2 norm or *sum of squared differences* SSD (2.6). If we denote the images to be compared I_1, I_2 we have:

$$\mathcal{D}^{SAD}(I_1, I_2) = \frac{1}{N} \sum_{(i,j) \in W} |I_1(i, j) - I_2(i, j)| \quad (2.5)$$

$$\mathcal{D}^{SSD}(I_1, I_2) = \frac{1}{N} \sum_{(i,j) \in W} \|I_1(i, j) - I_2(i, j)\|^2, \quad (2.6)$$

where N represents the number of pixels/voxels in the domain W .

SAD measure works better when the number of pixels/voxels is small and the intensities differences between images I_1, I_2 are large. Both measures can handle Gaussian noise.

In case the intensity values of one of the images is linearly shifted by different settings on the image acquisition scanner, or is affected by non-Gaussian noise distributions the *normalized cross correlation coefficient* (NCC) is introduced. NCC is given by:

$$\mathcal{D}^{NCC}(I_1, I_2) = \frac{\sum_{(i,j) \in W} I_1(i, j) I_2(i, j)}{\|I_1\| \|I_2\|} \quad (2.7)$$

In order for this measure to be effective in the registration process it must be computed on zero-mean images. The distance measure is called *zero-mean NCC* (ZNCC) and is expressed by:

$$\mathcal{D}^{ZNCC}(I_1, I_2) = \mathcal{D}^{NCC}(I_1 - \bar{I}_1, I_2 - \bar{I}_2), \quad (2.8)$$

where $\bar{I}_i = \mathbb{E}[I_i]$ is the expectation value of I_i , for $i = 1, 2$.

From a statistical point of view, SSD represents the *mean squared error* of the difference between the predictions I_1 and the true values I_2

While SAD and SSD are dissimilarity measures that should be minimized, NCC and ZNCC are measure of similarity between images, therefore they should be maximized. In fact, if we compute the SSD (or equivalently the mean squared error of the difference) on the normalized images $\tilde{I}_1 = \frac{I_1 - \mathbb{E}[I_1]}{\sigma(I_1)}$ and $\tilde{I}_2 = \frac{I_2 - \mathbb{E}[I_2]}{\sigma(I_2)}$ we have:

$$\begin{aligned}
\mathbb{E}[(\tilde{I}_1 - \tilde{I}_2)^2] &= \mathbb{E}[\tilde{I}_1^2 + \tilde{I}_2^2 - 2\tilde{I}_1\tilde{I}_2] \\
&= \mathbb{E}[\tilde{I}_1^2] + \mathbb{E}[\tilde{I}_2^2] - 2\mathbb{E}[\tilde{I}_1\tilde{I}_2] \\
&= 1 + 1 - 2\mathbb{E}\left[\frac{(I_1 - \mathbb{E}[I_1])(I_2 - \mathbb{E}[I_2])}{\sigma(I_1)\sigma(I_2)}\right] \\
&= 2 - 2\mathcal{D}^{ZNCC}(I_1, I_2)
\end{aligned}$$

From a geometric point of view, SAD and SSD may be understood as the norm of the vector that has as components the differences between images intensities, therefore is a measure to be minimized, meanwhile NCC and ZNCC can be seen as the cosine of the angle between the two vectors that have as components the intensities of each image, therefore is a measure to be maximized in order to achieve the perfect alignment.

Applications of these measures of distance in medical image registration may be found in [60] or [3].

A more recent distance measure for inter-modal images is the *residual complexity* (RC), proposed by Myronenko et al. [83]. This measure was conceived to register images corrupted by multiplicative non-stationary intensity distortions.

The choice of an appropriate distance measure is a harder task in the case of multi-modal imaging. There are two main approaches: the reduction of the multi-modal problem to a single-modal problem and the use of information theory to define the distance measure.

The first choice can be realized by deriving one modality from another, or by mapping both images in a common domain. Roche et al. [94] transformed a MRI image into a US image by exploiting the MRI intensities and MRI magnitude of the gradient in order to predict the US value. By exploring the physical properties of US, Wein et al. [123] simulated an US image from a CT image. To map the images in the same domain Maintz et al. [73] used morphological tools to create new gray-value intensity images. This method applied morphological opening and closing to extract edge information and then cross-correlation to align the images.

A compromise between SSD and mutual information, based on normalized image intensity gradients is introduced by Haber and Modersitzki in [41]. This distance measure is based on the observation that even for images of different modalities, intensity changes appear at corresponding positions. However, the gradient also measures the

strength of the change which is an unwanted information for multi-modal information, therefore the gradient is normalized by its norm. This *normalized gradient* of an image I is defined by:

$$NGF[I] = \frac{\nabla I}{\sqrt{|\nabla I|^2 + \eta^2}}, \quad (2.9)$$

where η is an edge parameter.

From here the distance measure is derived:

$$\mathcal{D}^{NGF}(I_1, I_2) = \int_w 1 - (NGF[I_1]^\top NGF[I_2])^2 dx \quad (2.10)$$

More recently, Heinrich et al. [48] introduced a new descriptor for multi-modal registration. The idea is the use of similarities between neighboring patches to define the descriptor. Once the descriptor is constructed the matching criterion is computed as a vector-difference.

The most important distance measure for multi-modal registration revolves around the so called *mutual information* (MI), and was popularized in image registration by two different groups: [124], [119] in US and [26], [69] in Europe. The MI is given by:

$$\mathcal{D}^{MI}(I_1, I_2) = H(I_1) + H(I_2) - H(I_1, I_2), \quad (2.11)$$

where $H(\cdot)$ is a measure of the uncertainty in a random variable and was first introduced by Shannon in [104]:

$$H[p(x)] = \mathbb{E}_x[\log \frac{1}{p(x)}] = - \sum_x p(x) \log p(x), \quad (2.12)$$

$H(\cdot, \cdot)$ is the joint entropy of I_1 and I_2 and measures the amount of information we have in the combined images. If the images are totally unrelated, then the joint entropy will be the sum of the entropies of the individual images, therefore \mathcal{D}^{MI} will be zero. The more similar (i.e., less independent) the images are, the lower the joint entropy compared to the sum of the individual entropies, therefore \mathcal{D}^{MI} will reach its maximum when the images are aligned.

The mutual information is the basis of further developments, starting from the work of Studholm et al. [111], where the normalized version of MI is introduced in order to overcome the dependence on the overlapping area of MI, or in [112] where regional mutual information (RMI) is defined, to Zhuang et al. which use d locally evaluated

MI in combination with global MI [131], Pluim et al. [91] which used intensity image gradient as an additional cue and so on.

2.2.2 Parametric and non parametric solutions

The solution of the registration process, or the registration transformation, may be generated from a physical model that constrains the registration by the smoothness \mathcal{S} term in (2.2), or by a parametrization of the transformation.

The type of the mapping is of paramount importance for the registration as it reflects the class of transformations that are desirable or acceptable, and therefore limits the solution to a large extent. The registration parameters estimated through the optimization step correspond to the degrees of freedom of the transformation involved. Their number varies greatly, from six in the case of global rigid transformations in the three dimensional space, to a number equal to the number of pixels/voxels of the image in the case of a dense transformation.

In the following, we shall call the transformations constrained to belong to a certain class of functions such as rigid, linear or affine, polynomial, radial basis functions, free form deformations, B-splines, thin plate spline, that have a relatively low number of parameters, *parametric transformations* and the transformations given as the discretized numerical solution of the (2.2), constrained by the chosen regularizer, *non parametric transformation*.

In the case of parametric transformations the models are derived from linear or non-linear interpolation or approximation theories. The non linear methods range from polynomial to spline-based transformations that are piecewise polynomial functions with a predefined degree of regularity (see [117] for details).

The most known models are the *radial basis functions* (RBF), where the value of an interpolation point x is given as a function of its distance r from the known sample p_i :

$$u(x) = \sum_{i=1}^n \beta_i \phi(\|x - p_i\|) \quad (2.13)$$

A comparative study for non rigid registration using RBF was published by Zagorchev et. al in [129].

Another popular parametric method is based on the *thin-plate splines* (TPS) functions, first introduced in image registration by Bookstein in [11].

The *free-form deformations* (FFD) use a rectangular grid $G = K_x \times K_y \times K_z$ superimposed to the image of the size $N_x \times N_y \times N_z$ that gets deformed under the influence of the control points. This approach comes from the computer graphics community [102], [51] and became popular in the medical image registration when coupled with cubic B-splines [30], [96], [63], [101].

FFD based on B-splines can be expressed as a three dimensional tensor product of one dimensional B-splines:

$$u(p) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^2 B_l(\mu_x) B_m(\mu_y) B_n(\mu_z) d_{i+l, j+m, k+n} \quad (2.14)$$

where $i = \lfloor x/N_x \rfloor - 1$, $j = \lfloor y/N_y \rfloor - 1$, $k = \lfloor z/N_z \rfloor - 1$, $\mu_x = x/N_x - \lfloor x/N_x \rfloor$, $\mu_y = y/N_y - \lfloor y/N_y \rfloor$ and $\mu_z = z/N_z - \lfloor z/N_z \rfloor$ and B_l correspond to the B-spline basis functions. The number of parameters varies with the number of control points of the grid G .

When the transformation is derived from physical models, the displacement is given as the reaction of the model to a force. The force is generated by the similarity between the images. The linear models or the *elastic body deformation*, first introduced by Broit [15] in medical image registration, is described by the Navier-Cauchy partial differential equation:

$$\mu \nabla^2 u + (\mu + \lambda) \nabla (\nabla \cdot u) + F = 0, \quad (2.15)$$

where F is the force field that drives the registration based on the matching criterion, μ refers to the rigidity and λ is Lamé's first coefficient.

To implement (2.15) in medical image registration the image grid is considered as an elastic membrane.

The *viscous fluid flow* approach in [22] models the image as a viscous fluid and the transformation is governed by the Navier-Stokes equation. These transformations are able to recover large deformations.

The *diffusion model* [28] is governed by the diffusion equation:

$$\Delta u + F = 0, \quad (2.16)$$

while in the *curvature registration scheme*, the deformation is modeled by the following equilibrium equation:

$$\Delta^2 u + F = 0. \quad (2.17)$$

This approach does not penalize affine linear transformations.

The *optical flow* techniques model the image as a function in space and time, $\mathcal{T} : W \times \mathbb{R}_+ \rightarrow \mathbb{R}$ and assume that a particle located at $x(t)$ at time t does not change intensity. Thus, $\mathcal{T}(x(t), t) = \text{const}$, and considering the velocity $v = \dot{x}$, we have:

$$\partial_t \mathcal{T} + \nabla_x \mathcal{T} v = 0, \quad (2.18)$$

which is the optical flow constraint [50], [6]. If we consider $\partial_t \mathcal{T} \approx \mathcal{T} - \mathcal{R}$ and we relax the condition in (2.18) to a minimization we may derive a similar formulation to the equation (2.2):

$$\mathcal{D}^{OF}[v] = \frac{1}{2} \|\partial_t \mathcal{T} + \nabla_x \mathcal{T} v\|^2 + S[v]. \quad (2.19)$$

The optical flow leads the deformable template toward the reference. Example of algorithms for medical image registration based on the optical flow are the original formulation of Thirion's demons algorithm [115], [116] and Christensen's fluid approach [21].

2.2.3 Optimization

The last step of the classical approach to image registration is the optimization. The aim of this step is to derive the optimal transformation that best aligns the two images according to the objective function given by the equation (2.2).

In this work we focus on the continuous optimization since most of the registration algorithms use this approach, even if there are approaches based on the discrete optimization.

The reasons for using a continuous model are because it is more practical, since the transformed object does not always align with the pixel/grid and because of its computational efficiency. To derive a continuous model from the discrete data, interpolation has to be used.

The major difficulties in image registration, from the optimization point of view, is the handling of a variety of local and even global minimum. Multilevel methods are thus essential and it is not recommended to solve a problem using one fixed level.

Except *Powell's conjugate direction method* [92], all the commonly used methods

are based on the computation of the gradient: *gradient descent*, *conjugate gradient*, *quasi Newton*, *Gauss-Newton*, *stochastic gradient descent*. A more detailed description of these and other methods of optimization may be found in [86].

When using an MI-based distance measure, one must pay particular attention to the optimization schemes that require the computation of the Hessian matrix due to its high computational cost. One of the most successful approach is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method which estimates the Hessian by an update from an initial approximation and a sequence of differences of search directions and gradients.

A comparative study of optimization strategies in image registration using mutual information is reported by Klein et al. in [61].

2.3 Experimental set-up for testing

To test some of the registration algorithms presented so far, we have used medical images obtained from a triple modality 3D abdominal phantom, developed by Computerized Imaging Reference Systems Inc (CIRS, Norwalk, VA). The model 057 (Figure 2.2) interventional 3D abdominal phantom, is designed to address minimally invasive procedures and to be used in different abdominal scan techniques such as CT, MRI and US (Figure 2.3), developing imaging protocol and system testing and validation .



Figure 2.2: CIRS 057 interventional 3D abdominal phantom.

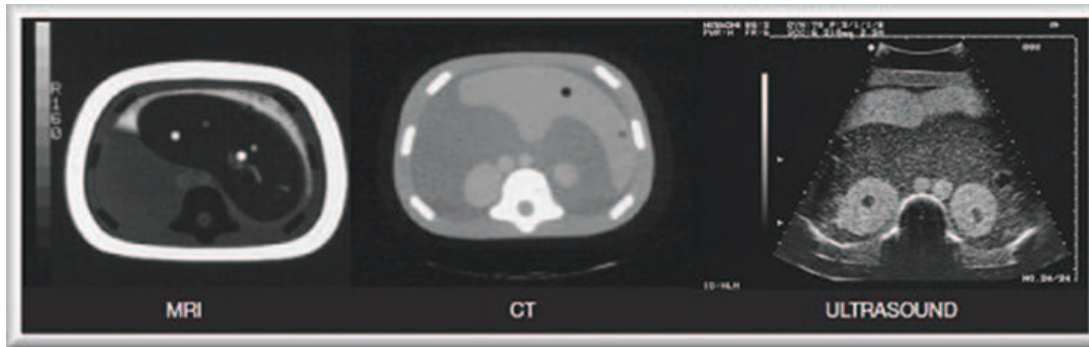


Figure 2.3: Different scannings of the phantom. From left to right: MRI, CT, US.

The reasons for the choice of this set-up, beside the multi-modal capability, is due also to the complete control over the physical position of the scanned sections. We wanted to obtain two completely aligned 2D slices in two different modalities, therefore to create the ideal framework for the registration algorithms, considering the nature of our problem which is the registration of real-time acquired images with accurate pre-operative scans. The difficulties encountered by the registration algorithms are due on one hand to the different physical process that generates the images and, on the other hand, by the deformation of the phantom generated by the pressure of the US probe and by the different covering of the US convex probe with respect to the CT slice which is a condition often encountered in practice (see Figure 2.4, the deformation takes place in the upper part of the US image).

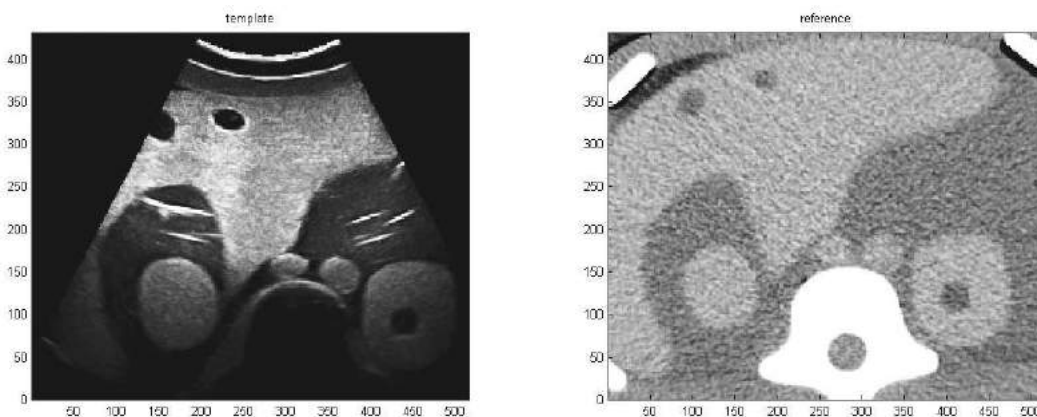


Figure 2.4: The images used in the registration process. On the left hand side the US template image; on the right hand side the CT reference (xed) image.

The Model 057 simulates the abdomen from approximately the thorax vertebra T9/T10 to the lumbar vertebra L2/L3 using simplified anthropomorphic geometry.

The primary organs included are the liver, a portion of the lung that surrounds the liver, portions of the portal vein, abdominal aorta and inferior vena cava, and partial kidneys (Figure 2.5). Embedded within the liver are simulated lesions. The simple transverse ribbing, simulated spine, urethane membrane and ABS end-caps provide framing and protection, making the Model 057 durable enough for extended scanning sessions yet enabling insertion of various surgical instruments as needed.

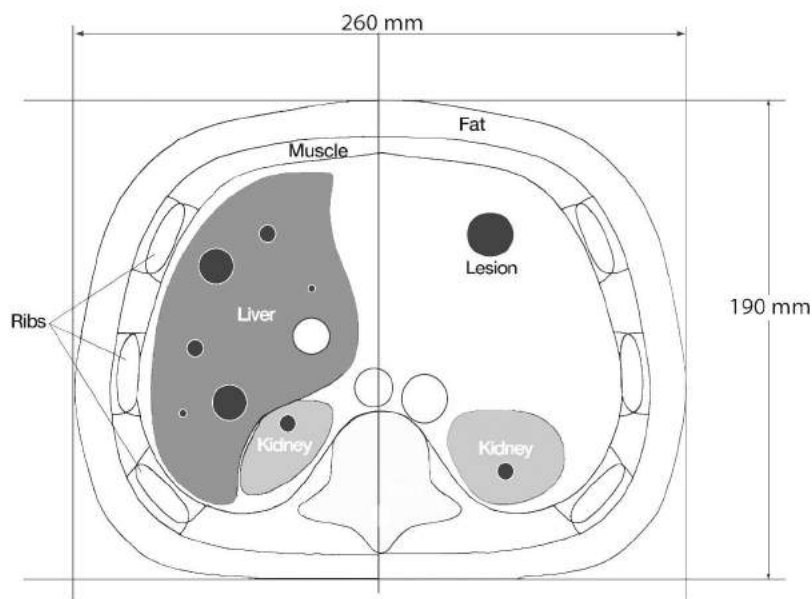


Figure 2.5: Internal anatomy of the anatomical phantom.

In minimally invasive procedures the main interest of the medical community is to register pre-operative accurate images with intra-operative real-time acquired images, therefore the modalities of acquisition we have used here are CT for the pre-operative stage and US for the intra-operative stage (Figure 2.4).

The phantom was equipped with 4 markers in order to register it with the CT dataset (Figure 2.6). The global coordinate system is given by an optical tracking system composed by infrared light emitting cameras.

The US images were acquired with a 2D probe equipped with markers in order to map its position to the global coordinate system. The calibration process of the US probe converts the point (u_i, v_i) in image plane coordinates (pixels) in 3D homogeneous coordinates defined with respect to the global reference system by the following formula:

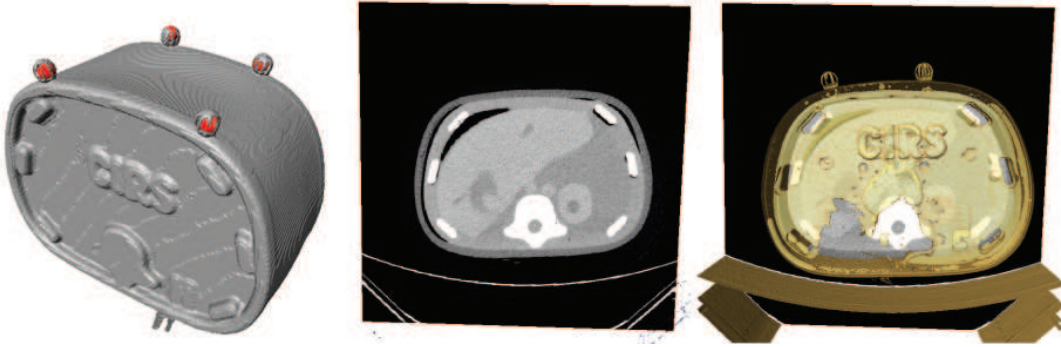


Figure 2.6: From left to right: the 3D image of the phantom with 4 markers that identify its position in space; a CT slice of the phantom; the same CT slice with the outline of the phantom in transparency.

$$\begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = T_p T_{pi} \begin{pmatrix} s_u u_i \\ s_v v_i \\ 0 \\ 1 \end{pmatrix}, \quad (2.20)$$

where T_p is the pose matrix that encodes the pose of the markers and is given by the tracking system, while T_{pi} is the transform estimated by the calibration procedure together with the scale factors s_u and s_v that we assume isotropic, therefore $s_u = s_v = s_{uv}$ (Figure 2.7).

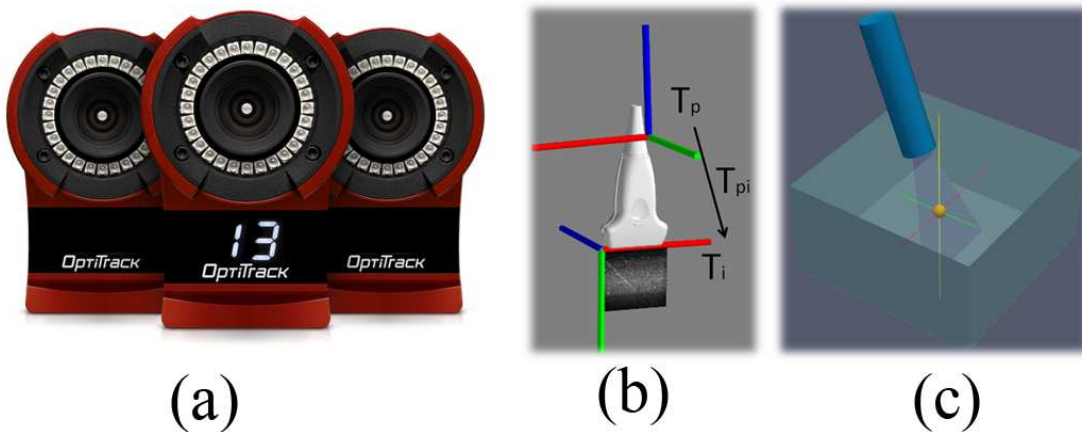


Figure 2.7: (a) The optical tracking system defines the global reference system. (b) The US probe and the transformations involved in the calibration process. (c) The calibration of the US probe is based on the identification of a fixed point in different poses of the probe.

We have used methods for multi-modal registration with transformation ranging from parametric (affine, spline) to non-parametric (based on curvature or elastic) and different optimization methods. In the case of non-parametric registration the multi-level approach was necessary.

Appendix A reports the visual and numerical results of the registration algorithm we have tested.

Even in this controlled set-up with a strong prior given by the initial registration, the quality of the results we have obtained is very bad, making the application of these methods in the OR scenario completely useless.

It is well known that there is no best algorithm for registration but it seems to us that each algorithm was designed to solve the registration of a certain dataset and it makes no sense to use it on different datasets. This may be the reason of the existence of such a large number of registration algorithms.

The next section reports the drawbacks we found during the tests.

2.4 Why the medical image registration framework cannot work for OR applications?

The major issues we have identified that makes this approach unfeasible in OR applications are:

- The distance measure MI: this measure is very sensible to noise, incomplete data, no completely overlapping domains.
- The distance measure NGF: in our tests this measure has not produced the results we expected.
- The optimization: since the objective function is highly non convex, all the optimization methods fail in finding the global minimum or maximum.
- The multilevel approach: even though the multilevel approach is vital when using the non-parametric approach, in the case of multi-modal images the low level approximation of the registration transformation is not accurate at all and generates wrong results at higher levels.
- The computation time is very high even when only a couple of 2D images is involved.

Hadamard [42] defined a problem well-posed if it has a solution, this solution is unique and depends continuously on the data. In this sense, the registration problem is ill-posed since for every spatial location $x \in W \subset \mathbb{R}^d$, we search a vector $f(x) \in \mathbb{R}^d$, but usually only a scalar information $\mathcal{T}(y(x))$ is given.

[95] gives a simple example of how ill-posedness may give arbitrary results that seem very good from an optimization point of view but are completely useless in practice. The authors called this approach CURT (completely useless registration tool), which is a very simple registration algorithm based on correspondence of pixels sorted by increasing intensities, and showed that this method outperforms other registration methods in the following cases: SSD difference, NCC image correlation and NMI (normalized MI) image similarity.

Except the well established and accepted solutions for image registration, some newer solutions were introduced and most of them address new distance measures such as a locally evaluation of MI in combination with standard global MI [131], the residual complexity to account for complex spatially varying intensity distortions [83], learning based multi-modal registration using Kullback-Leibler divergence for non-rigid registration [39] and rigid registration using learning based Jensen-Shannon divergence [65].

These new methods tend to ultimately complicate the already complex registration framework and most of them follow the three steps approach.

It would be difficult to test all of them but we intuit that our case study cannot be solved by the classical approach of medical registration framework and we shall need additional information to solve the registration in the OR.

Except for the rigid transformation most of the solutions yield a transformation that has no physical meaning, therefore even if one of the solutions would function, it becomes difficult to validate the results for OR applications.

The next chapter introduces the landmark-based approach, which is a simplified method for image registration based on very sparse data. After the introduction of the approach, a simple 3D registration algorithm for the OR set-up is described.

Chapter 3

Registration of binary images

Always try the problem that matters
most to you.

Andrew Wiles

In chapter 2 we have analyzed and tested the most significant algorithms for medical image registration and we have underlined that none of them would be suitable for an application in the OR that requires the integration of images coming from different modalities. The only advantages of such algorithms are that they are automatic, once some parameters are fixed, and they do not need prior constraints or further interactions with the users.

A different approach in image registration, that we investigate here, is the so called feature-based or landmark-based registration. These methods rely on the identification and matching of corresponding points. The points or landmarks are usually placed in salient image locations, which are considered to correspond to meaningful anatomical regions. The solution of the registration is obtained in a straightforward manner once the landmarks are individuated. The bottleneck in this case is the location of the landmarks and the establishment of the correspondences among them.

3.1 Introduction

In the case of landmark registration the input of the algorithm is represented by two sets of landmarks. The first set of landmarks include points belonging to the template image and the second set is composed by landmarks from the reference image (see Figure 3.1).

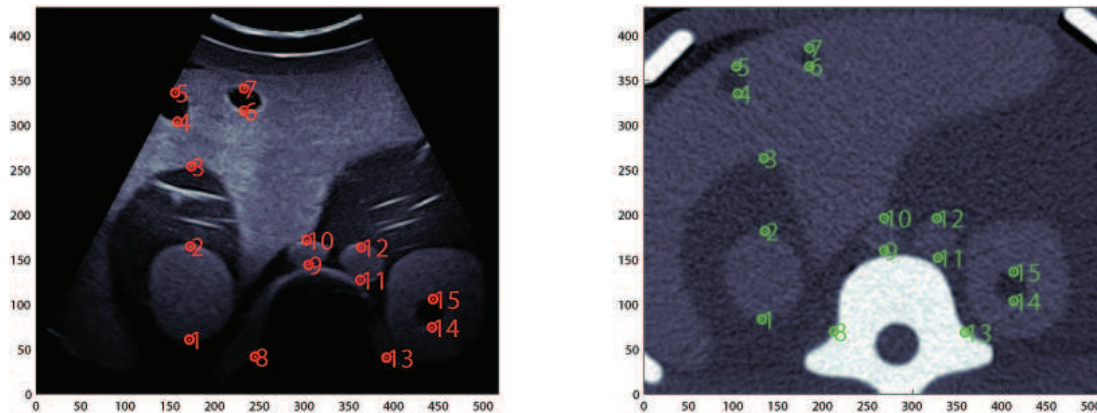


Figure 3.1: The template and the reference images with corresponding landmarks.

The algorithm must solve the correspondences and the transformation. Even if there are approaches that solve only the correspondences, they are used together with interpolation to find dense correspondences between the images.

3.1.1 Extraction of landmarks and binary images

The main limitation of the medical images is that they are not as rich in details as for instance the digital photographs. The extraction of landmarks has been studied more in the case of 2D images and less in the case of 3D images.

In the computer vision literature there is a huge amount of work dealing with point of interest extraction. Harris et al. [47] proposed a combined corner and edge detector based on the local auto-correlation function and many other works extended this approach. For instance, Shi and Tomasi [107] used the minimal eigenvalue of the structure tensor in order to track point of interest.

Another significant work by Lowe [67] introduced the scale invariant feature transform (SIFT) algorithm to extract significant points by using the difference of Gaussians function applied in scale space to a series of smoothed and resampled images. Many variants of SIFT have been proposed such as PCA-SIFT [56], affine SIFT [81] or SIFT in higher dimensions [20] used also in the context of medical imaging by matching landmarks between 3D MRI images and 3D CT images changing over the time (4D CT images). The speeded up robust features (SURF) algorithm [8] is partially inspired from SIFT but with an increased efficiency.

A method used most in medical image registration is based on fiducial markers and one of the main points of interest regards the correlation of the errors between the

ducial localization error (FLE) and the target registration error (TRE), especially in image guided interventions, [34], [126], [80].

In medical images, boundaries or surfaces are frequently more distinct than landmarks, and various segmentation algorithms can successfully locate high contrast surfaces. This is especially true for skin surface. If equivalent surfaces can be automatically segmented from two images to be combined, then registration can be achieved by fitting the surfaces together.

In this case, the image to be registered is a processed version of the original data defined by (2.1), that we call it *binary image*, since its codomain is the set $\{0, 1\}$:

$$\mathcal{T} : W \rightarrow \{0, 1\}, W \subset \mathbb{R}^d, \quad (3.1)$$

where $\mathcal{T}(x) = 1$ if the pixel/voxel x from the image domain W represents a point on the extracted surface, and $\mathcal{T}(x) = 0$ otherwise.

By introducing this definition, some of the notions from chapter 2 can be reused.

3.1.2 Correspondences and transformation

When the correspondences between landmarks are given, *Procrustes analysis* is a popular method for shape analysis, see for instance [27]. This technique is the departing point of our *generalized shape* theory given in the chapter 4, and the previous solutions will be discussed there.

The transformation can be estimated using interpolation strategies when the correspondences are known (see 2.2.2). Some results using this approach will be presented in the section 3.2.

The most well-known method that infer both the correspondences and the transformation is the iterative closest point (ICP) method, proposed by Besl and McKay [10]. This algorithm is designed to work with different representations of surface data: point sets, line segment sets (polylines), implicit surface, parametric curves, triangle sets, implicit surfaces and parametric surfaces. For medical image registration the most relevant representations are likely to be point sets and triangle sets, as algorithms for delineating these features from medical images are widely available.

The algorithm is able to register a data shape P with N_p points to a model shape X with N_x primitives. For each element of P , the algorithm first identify the closest point on X shape, then finds the least square rigid-body transformation relating these pairs of point sets. The algorithm then redetermines the closest point set and continues

until finds the local minimum match between the two surfaces, as determined by some tolerance threshold.

If the original representation of the binary data surface P is converted to a set of points $\{p_i\}$, the first stage identifies, for each point p_i the closest point on the model surface X . This is the point $x \in X$ for which the distance d between p_i and x is minimum.

$$d(p_i, X) = \min_{x \in X} \|x - p_i\| \quad (3.2)$$

The resulting set of closest points (one for each p_i) is $\{q_i\}$. For a triangulated surface, which is the most likely model representation from medical image data as for instance CT, the model X comprises a set of triangles t_i . The closest model point to each data point is found by linearly interpolating across the facets. If triangle t_i has vertices r_1, r_2, r_3 , then the smallest distance between the point p_i and the triangle t_i is

$$d(p_i, t_i) = \min_{u+v+w=1} \|ur_1 + vr_2 + wr_3 - p_i\| \quad (3.3)$$

where $u, v, w \in [0, 1]$. The closest model point to the data point p_i is, therefore, $q_i = (ur_1, vr_2, wr_3)$.

A least squares registration between the points $\{p_i\}$ and $\{q_i\}$ is then carried out using the solution to the Procrustes problem [31]. The set of data points $\{p_i\}$ is then transformed to $\{p'_i\}$ using the calculated transformation, and then the closest points are once again identified. The algorithm terminates when the change in mean square error between iterations falls below a defined threshold.

The optimization can be accelerated by keeping track of the solutions at each iteration. As the algorithm iterates to the local minimum closest to the starting position, it may not find the correct match. The solution proposed in [10] is to start the algorithm multiple times, each with a different estimate of the rotation alignment, and choose the minimum of the minimum obtained.

In most applications not every point has a corresponding match, so a simple distance threshold can be used to discard correspondences in order to deal with partially overlapping scenes. Additionally, to further improve the quality of correspondences, a lot of efforts have gone into the area of feature selection [106], as well as including extra information such as colors [54] or normals [4] that could improve the correspondence problem. Since ICP requires an exhaustive search through the correspondence space, several variants that address the problem of its computational complexity have

been proposed [87], [97].

The main limitation of the ICP algorithm is due to the fact that it can handle only small movements of the two point clouds to be registered, therefore is a local method. To avoid this limitation, Gelfand [35] proposed the use of an integral descriptor [74] based on local geometry invariant to rotations, translations and robust to noise. Only a small number of feature points are picked from the data shape according to the uniqueness of the descriptor at the point. For each feature point of the data, the corresponding point of the model is chosen using the descriptor values. A similar approach, but with the use of different signatures for each point, is presented by Rusu et al. in [98].

A new approach, proposed by Mitra et. al in [76], for the local registration of point clouds, uses the time coordinate for each acquisition and pose the problem of computing the motion of each frame as an estimation of certain kinematic properties of the resulting space-time surface.

The so called 'boot-strapping' or the initialization of the ICP algorithm with the use of high dimensional feature descriptors such as SIFT and SURF is introduced by Pandey et al. in [89]. In [33] the initial rigid body transformation is computed after the identification of the correspondences in the high dimensional feature space in a random sample consensus (RANSAC).

A generalized-ICP algorithm that combines ICP with 'point to plane ICP' into a probabilistic framework is presented in [103].

3.2 Test on phantom data

We compare here the landmark-based algorithms performances with the intensity-based algorithms, that we have seen in the previous chapter. It is obvious that, in the case of landmark-based algorithms. only sparse information is used, while in the case analyzed in the previous chapter all the radiometric information is used.

We assumed that the correspondences are known *a priori* and the landmark-based registration techniques we test are linear, polynomial and quadratic.

If we denote $\{T_i\}$ the set of landmarks in the template image and $\{R_i\}$ the set of landmarks in the reference image, the goal is to find a transformation f such that

$$f(T_i) = R_i, \forall i. \quad (3.4)$$

Observe that this approach, called Lagrangian, is the opposite of the Eulerian approach presented in the chapter 2, equation (2.3). The same framework used in chapter

2 may be implemented by simply substitute f with f^{-1} , when feasible.

More generally, the interpolation conditions (3.4) are replaced with approximation conditions:

$$\mathcal{D}^{LM}(f(\mathcal{T}), \mathcal{R}) = \sum_i \|f(T_i) - R_i\|^2 = \min. \quad (3.5)$$

If we consider the images to be registered given in the binary form (3.1), this is nothing more than the SSD measure.

The linear solution, when the image domain is a subset of \mathbb{R}^d , looks for a $d \times d$ matrix A and a d -dimensional vector v such that $f(T_i) = AT_i + v$ and the minimization in the equation (3.5) takes place. Given n landmarks, the number of conditions is nd , while the number of parameters is $d(d+1)$. Considering the low dimensionality of the data and a relatively high number of landmarks, the match is not perfect, except in the case when $nd = d(d+1)$ (see appendix B).

Enlarging the transformation space, for example by choosing polynomial transformation, the number of parameters increases with the polynomial degree. If we write the transformation f in vectorial form $f = (f_1, \dots, f_d)$ and $T_i = (T_i^1, \dots, T_i^d)$ and we consider the quadratic transformation, we have:

$$f_j(T_i) = a^j + \sum_{k=1}^d b_k^j T_i^k + \sum_{m \neq n} c_{m,n}^j T_i^m T_i^n + \sum_{k=1}^d (T_i^k)^2 \quad (3.6)$$

From (3.6) it follows that the number of parameters of the quadratic transformation is $\frac{d^2}{2} + \frac{3}{2}d + 1$.

If we want to solve (3.4) and find a function whose bending energy is minimum, we obtain the thin-plate-spline (TPS) transformation [121]. The bending energy is the integral of the square of the second derivative:

$$\mathcal{S}(f) = \sum_{j=1}^d \int \sum_{m,n=1}^d \left(\frac{\partial^2 f_j}{\partial x_m \partial x_n} \right)^2 dx_1 \dots dx_d, \quad (3.7)$$

where $f = (f_1, \dots, f_d) : \mathbb{R}^d \rightarrow \mathbb{R}^d$.

Therefore the equation we want to solve is

$$\mathcal{S}(f) = \min, \quad (3.8)$$

subject to (3.4).

The solution of this equation belongs to a space that is spanned by shifts of a known

radial basis function ρ by a linear term:

$$f_j(x) = \sum_i c_j^i \rho(\|x - T_i\|) + Ax, \quad (3.9)$$

where A is a $d \times d$ matrix representing the affine transformation, $c = \{c_j^i\}$ is a $n \times d$ warping coefficient matrix representing the non-affine deformation and $\rho(r) = r^2 \log(r)$ represent the kernel function.

The interpolation condition (3.4) can be relaxed and replaced by an approximation condition if we solve for both (3.5) and (3.8):

$$\mathcal{D}^{LM}(f(\mathcal{T}), \mathcal{R}) + \alpha \mathcal{S}(f) = \min, \quad (3.10)$$

where $\alpha \geq 0$ is a parameter that balances the interpolation and the approximation scheme. This equation is nothing more than the registration formula we have already introduced in chapter 2 (see equation (2.2)). The solution is again given by (3.9) [78].

A nice property of the TPS is that it can always be decomposed into a global affine and a local non-affine component. Consequently, the TPS smoothness term in (3.10) is solely dependent on the non-affine components. This is a desirable property, especially when compared to other splines, since the global pose parameters included in the affine transformation are not penalized.

Appendix B reports the registration results. The set-up and the images are the same ones used in the previous chapter.

3.3 Discussion

3.3.1 Why the binary (landmark) registration works better than the intensity-based registration for an OR application?

The first observation is that, in both cases, the transformations we are interested in have no physical meaning. Except in the rigid case, they are just mathematical tools that help to solve the registration. Following this observation, the intensity based completely automatic algorithms fail and give arbitrary results if no prior information is given. In the case of binary registration, even with a small number of landmarks and the rigid/linear constraint the results are much better. Extracting corresponding landmarks in multi-modality images excludes the computation of a multi-modal distance measure. As we have seen in the previous chapter, the computation of MI and NGF generates

misleading results and the computational time is very high, even in the 2D case. On the other hand, the binary registration will always yield a decent result and the computation time is very low.

3.3.2 What are the limitations of the binary registration for an OR application?

Even if the registration results in the binary case are preferable, the bottleneck of this method is not only the detection of landmarks, but also the computation of the correspondences. Having the correspondences, we can choose a mathematic model and, based on the number of landmarks, the registration is solved as some form of interpolation (section 3.2). This solution is simple and intuitive and it may always isolate a linear transformation, even with the TPS solution, which is an advantage since in the OR scenario a rigid motion is always required.

3.3.3 What is the direction in which we should proceed?

We may consider the input data as given in a binary form. There is a large amount of work in the segmentation of medical images and the scope of this thesis is not about it. The binary images allow us to keep all the relevant information and the number of extracted features can range from some sparse data to dense information. Most of the radiologic information will disappear but, as we have seen, we don't need all the radiologic information. In fact, the radiologic information should be used only to detect the relevant features.

The rigid transformation is always needed as a global alignment and, when required, a non-rigid transformation should be derived as a local re nement.

3.4 A real-time 3D set-up in the OR

After these considerations, we searched for a registration algorithm that does not need the correspondences to find even a simple transform and the transform we shall

Some parts of 3.4 are based upon: B. Maris, D.Dall'Alba, P. Fiorini, A. Ristolainen, L. Li, Y. Gavshin, A. Barsi, V. K. Adhikarla *A phantom study for the validation of a surgical navigation system based on real-time segmentation and registration methods*, International Journal of Computer Assisted Radiology and Surgery, IJCARS (submitted).

always need in the OR is the rigid transform. In some cases a small deformation may be allowed.

The motivation for the implementation of a rigid solution, beside the computational cost, is that most of the (successful) systems employed currently in the OR, using the registration as a component, are based on the rigid solution. We recall here some of these devices:

- *Artemis* is a 3D ultrasound-guided prostate biopsy platform produced by Eigen (Grass Valley, CA, US) that uses the fusion of MRI for in-bore biopsies. During biopsy in the urology suite, *Artemis* with ProFuse Box helps target the suspected lesions (identified in ProFuse) using its real-time ultrasound guidance. This device integrates a registration module based on the manual identification of landmarks followed by TPS interpolation (see [40]) and is validated using a custom-built phantom.
- The *Explorer* navigation system from Pathfinder Technologies (Nashville, TN, US) designed to increase precision and physician confidence by providing advanced 3D visualization of anatomic structures and real-time tracking of therapeutic delivery tools in liver interventions: partial hepatectomy, extended hepatectomy, liver tumor ablation, living donor liver transplantation. The registration module uses as input the CT scan of the liver and a range scanner of the liver surface acquired during the intervention. The physician acquires manually some predefined landmarks using a tracked tool for the initial registration, then the input data together with the ICP algorithm are used to further refine the registration (see [25]).
- *Cas-One*, produced by Cascination (Bern, Swiss), is a stereotactic accessory for instrument guidance in open liver surgery. It visualizes in real-time the position and pose of surgical instruments relative to a three-dimensional model of the patient's liver. The registration is done by the manual selection of landmarks, assuming a locally rigid and temporarily static scenario [90]. After deformations occurring during the procedure, efficient means for registration updates are provided. The assumption of locally rigid patient registration was validated with nine clinical cases.

Some of these algorithms require the correspondences of landmarks and some of them are based on ICP algorithm, or variants of it, to perform the global alignment

rigidly. The main limitations in the OR is the difficulty to obtain correspondences on one hand, and the need of initialization of the transformation to avoid local minimum and the fact that the error threshold and the number of iterations must be limited beforehand, when using the ICP algorithm, on the other hand.

The algorithm we are proposing to solve the registration problem uses the binary representation (segmentation) of the images as given by the equation (3.1). The intensity information serves, in this case, only on the pre-processing (segmentation) phase. The binary image may be considered as a density function or mass distribution. From this distribution we derive the principal-axes or the principal moments of the image.

The principal-axes algorithm allows fast registration of large amount of segmented data. As we shall see later on, the problem may be reduced to the eigen-decomposition of two 3×3 matrices and there are very few registration parameters making it suitable for the real-time registration of 3D datasets. The intra-operative segmentation process produces binary images given by their characteristic function (equation (3.1)).

3.4.1 Realistic phantom set-up

We have evaluated the registration in an in-vitro set-up especially designed for our purposes. A plastic phantom with realistic shape and radiological properties was built (Figure 3.2).



Figure 3.2: The CAD model for the realization of the phantom (left) and the pancreas phantom (right) with the internal duct.

The design of a patient specific organ phantom begins by segmenting the CT scans of the patient abdomen. The segmentation is realized with 3D Slicer software, which

also gives the possibility to create a 3D representation of the organs surface. From this surface the model of the negative mold for model casting was created.

In the phantom development gelatin gels were used as they are self-supportive solids with a proper range of achievable elasticity, US characteristics and a relatively simple manufacturing process. The gel consists of food gelatin and distilled water mixed with formaldehyde and graphite flakes. The material fabrication has been well reported in different works, e.g. [68], [18], [45]. To mimic the internal structure a pipe was added as the pancreatic duct with a diameter of 4 mm and two cysts with a diameter of 20mm (Figure 3.2). In the preparation of the cysts no graphite was added to the gelatin mixture (to achieve low US attenuation). Five similar US capable pancreas phantom were built with pathology by using the same technique. The casted and hardened pancreas phantom was later inserted into a gelatin medium to make the phantom usable for US imaging with a linear probe. The pipe was removed after the surrounding gelatin hardened and the empty space left from the pipe was filled with water. The Hounsfield units and US attenuation values of the phantoms were in the same range of real tissue. The physical properties were not re-tuned in this study while the phantom was used to study the real-time registration algorithm.

3.4.2 Real-time US 3D acquisition and image segmentation

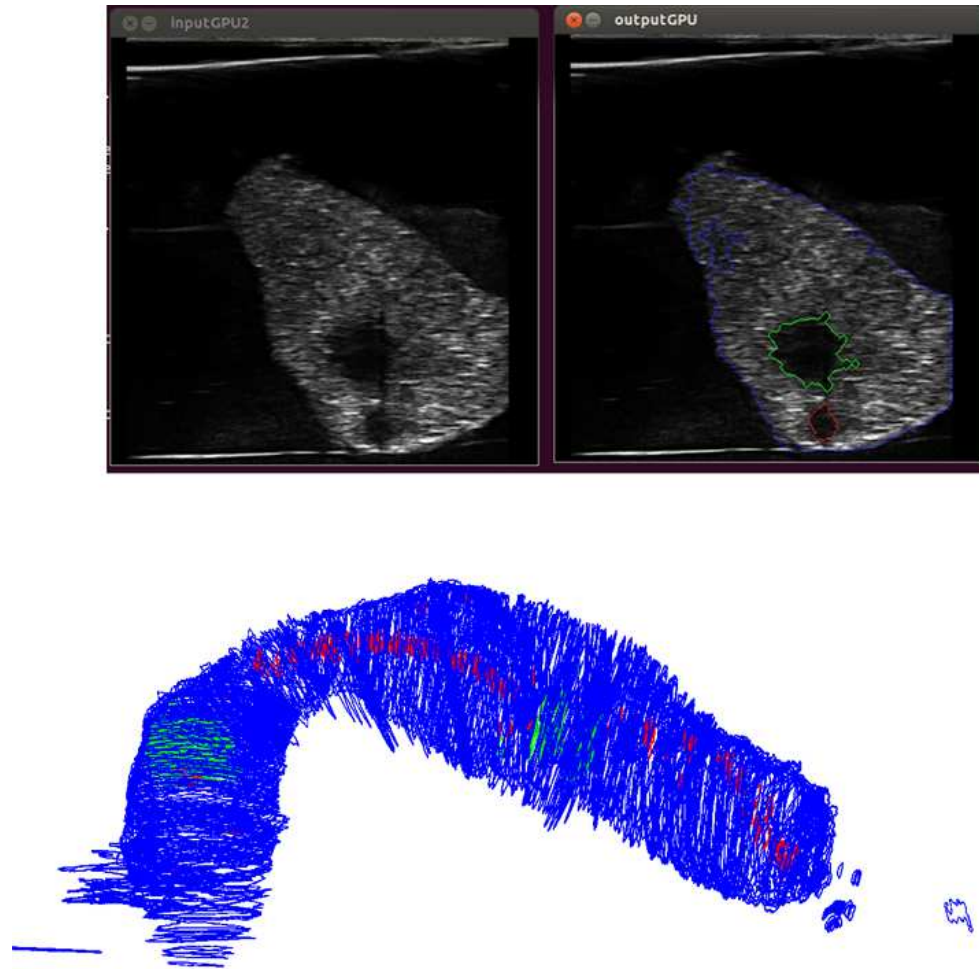


Figure 3.3: Top: A US slice before (left) and after the segmentation process (right), where the blue line is the organ surface, the green line is cyst's edge and the red line is edge of a duct. Bottom: 3D reconstruction of the segmented surface (blue) with duct (red) and cyst (green).

For the US acquisition and calibration processes please refer to 2.3.

Thanks to the calibration step we could process single 2D images (Figure 3.3), and the results of the segmentation could be reconstructed in 3D. These synthetic US images allow us to use simple segmentation methods.

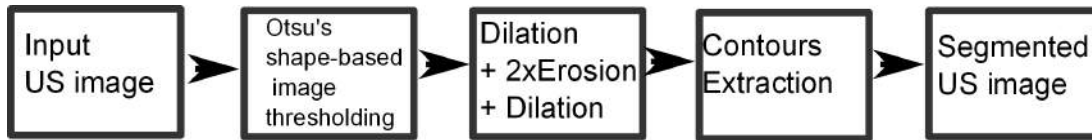


Figure 3.4: Real time US image segmentation pipeline.

Although the region growing approach [46] can be considered for segmentation, we might need a seed pixel location that must be provided as an input parameter. On the other hand, the region based approaches often need the output information from the previous processed pixel in order to process the current pixel. This may not be favorable in case of the process parallelization, which is very much required for real-time performance.

The main idea is to cut-off unwanted pixels by thresholding and then extract the cyst contour by calculating the gradient information. The proposed method *rst* aims to isolate the organ region from the background noise and then tries to extract required contour areas of the cyst. The *rst* step in this process is cut-off the unwanted noisy pixels from the ultrasound image. This can be done by using Otsu's method [88], which is an automatic thresholding method able to find the threshold value that minimizes the intra-class variance while maximizes the inter-class variance. When using continuous sequences of incoming images in the OR it may not be possible to provide the required threshold value manually, and then the automatic method used solve this problem.

A morphological dilation was used to fill the small holes inside the US image and to ensure an edge preserving expansion. A diamond shaped structuring element is used with a dimension of 7×7 pixels. The same structuring element is used for the subsequent double erosion step, followed by a second dilation step. All these morphological operation are necessary to retain the original contour (both external organ surface and internal structures) while removing false identified edges. Although the edge noise from the image is limited by using Otsu method, the output image still contains some speckles, which needs to be processed to get the final segmented US image. This is done in the contour extraction step. As we are working on the binary image, a simple gradient based contour extraction gives us all the possible contours on it. The extracted contours are then analyzed to measure how circular is the shape of a region by calculating the ratio of furthest to closest distance of each contour point to the contour's centre of gravity. As in an US image, the organ appears to be the largest object and the largest among the available contours can be directly considered to be the organ contour. The extracted contour of the organ is filled and a simple

bitwise AND operation is performed on the resulting image using the image obtained from morphological operations. This leaves us an image that has ones, just in the area where we have the cyst or a blood vessel (Figure 3.3). We can then follow the same gradient based contour extraction and this time, the output is directly the cyst contour.

The overall process chain of the proposed US segmentation algorithm is shown in Figure 3.4.

3.4.3 Real-time surface registration

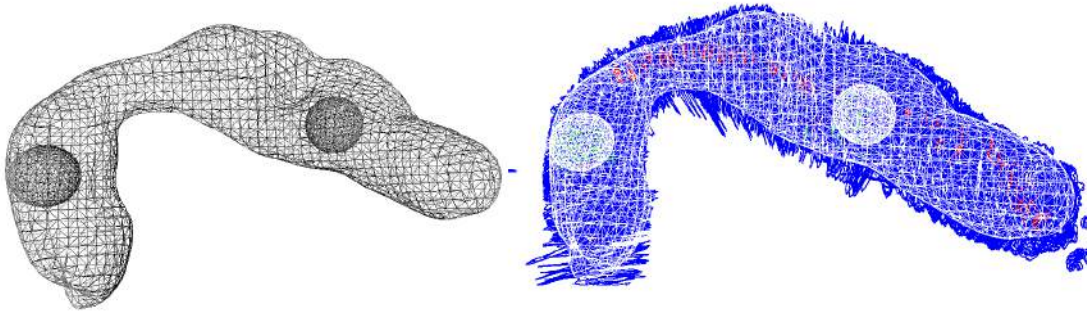


Figure 3.5: Left: CAD extracted surface. Right: US segmented image (blue) registered with the CAD surface (white).

The principal-axes registration was first introduced by Alpert et. al in [1]. We follow here the description of the algorithm given by Modersitzki in [77] with the difference that we use three-dimensional binary images as input.

The pre-operative binary images can be extracted from the original CT scan or from the CAD surface used in the phantom construction (Figure 3.2 and 3.5).

Having the real-time segmentation, we may impose the constraint that the deformations are small because the model is updated continuously. In this case, the binary images have the same distribution along the principal-axes and the deformations will be computed after the alignment along the principal-axes.

We define, for a binary image $I : W \in \mathbb{R}^d \rightarrow \mathbb{R}$, the expectation value of a function f with respect to I , by:

$$\mathbb{E}_I [f] = \frac{\int_{\mathbb{R}^d} f(x)I(x)dx}{\int_{\mathbb{R}^d} I(x)dx} = \frac{\int_W f(x)dx}{\int_W I(x)dx} \quad (3.11)$$

The center of an image I is defined by $c_I = \mathbb{E}_I[x] \in \mathbb{R}^d$ and the covariance by $Cov_I = \mathbb{E}_I[(x - c_I)(x - c_I)^\top] \in \mathbb{R}^{d \times d}$.

The center and an eigen-decomposition of the covariance matrix are used as features for the image. The resulting registration is named principal axes transformation.

Given two 3D binary images \mathcal{T} and \mathcal{R} with centers $c_{\mathcal{T}}$ and $c_{\mathcal{R}}$ and non-singular covariance matrices $Cov_{\mathcal{T}}$ and $Cov_{\mathcal{R}}$, after the transformation $\mathcal{T}(x) := \tilde{\mathcal{T}}(Rx + t)$ we have:

$$c_{\tilde{\mathcal{T}}} = c_{\mathcal{R}}, Cov_{\tilde{\mathcal{T}}} = Cov_{\mathcal{R}}, \quad (3.12)$$

where $R = D_{\mathcal{T}}\Sigma_{\mathcal{T}}U\Sigma_{\mathcal{R}}^{-1}$ is a 3×3 orthonormal matrix, $t = c_{\mathcal{T}} - Rc_{\mathcal{R}} \in \mathbb{R}^3$, U is a 3×3 unitary matrix, and the eigendecompositions

$$Cov_{\mathcal{R}} = D_{\mathcal{R}}\Sigma_{\mathcal{R}}^2D_{\mathcal{R}}^\top, Cov_{\mathcal{T}} = D_{\mathcal{T}}\Sigma_{\mathcal{T}}^2D_{\mathcal{T}}^\top$$

are used.

The solution is given again following the Eulerian approach (equation (2.3)), as opposite to the direct transformation of the tissue points, as in the case of the Lagrangian approach (equation (3.4)).

For normalization purposes we arrange the columns of $D_{\mathcal{R}}$ and $D_{\mathcal{T}}$ sorting the eigenvectors according to the magnitude of their eigenvalues. It is also possible that two or even three eigenvalues will be equal. These cases occur when the images A and B contains symmetries. In this case it is difficult to choose the axes. However, this situation almost never occurs for the images in real-life application, as in our tests.

3.4.4 Experimental test and results

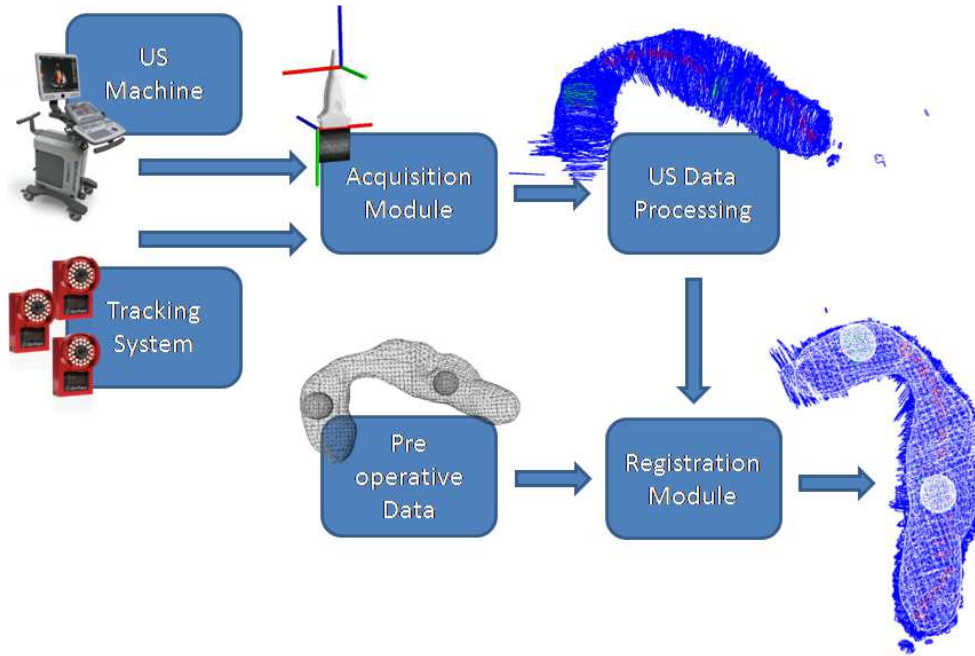


Figure 3.6: Block diagram of the proposed framework

We have evaluated the navigation system on 5 different phantoms and, for each of them, 10 acquisition tests were performed. The US device used in our setup is a Sonix MDP (Ultrasonix, Richmond, CA, US) that supports digital data acquisition. The pose of the probe is measured with an Optitrack (Natural Point, Corvallis, OR, US) passive IR optical tracking system. The calibration of the probe is performed with a single cross wire phantom, as described in the previous chapter (see section 2.3).

Figure 3.6 presents all the elements composing our system.

The US segmentation algorithm is implemented using OpenCV 2.4.1 with CUDA 4.0 and runs on a computer with Intel i7-920 CPU and NVIDIA Geforce GTX260 GPU. We tested the proposed algorithm on US images of the pancreas phantom with two cysts and a duct inside. Figure 3.4 shows the result of the proposed algorithm, the blue line is the organ surface, the green line is cyst's edge and the red line is edge of a duct. Each edge point is reconstructed in 3D thanks to the calibration transformation described before.

To evaluate the navigation system we have first computed the global registration error (GRE) as the RMS of the closest point distances of the segmented points on the

organ surface and the reference surface which is the CAD model used in the phantom manufacturing. The global error is given by the sum of the tracking, calibration, segmentation, registration errors, as well as the phantom construction error.

Along with the global registration error, we have computed the target registration error (TRE), which is given by distance between the centroid of the tumor in the CAD image and the same centroid in the segmented image. The TRE is the true measure of registration accuracy of importance to the success of a therapeutic procedures, and should be the only measure to be considered for the registration quality.

For the real time performance we have tested separately the segmentation algorithm on acquired images (size 408×408), and the time for the whole algorithm is 0.019 seconds. Thus, the algorithm supports real-time image segmentation at around 53 frames per second.

Phantom	GRE (mean/STD) mm	TRE (mean/STD) mm	Time (mean) sec
1	5.5207/1.8812	3.8707/1.5435	0.9654
2	6.0236/1.9212	3.8703/1.2371	0.9543
3	6.0045/2.0425	4.0030/1.8422	1.0104
4	5.9842/2.1203	4.0012/1.5809	0.9901
5	6.0103/1.9032	4.1302/1.7680	0.9882

Table 3.1: Accuracy and the precision of the system, both for the GRE and TRE, as the average and the standard deviation of the error. The last column gives the average time taken by the entire system to accomplish the registration.

The error obtained in the GRE and TRE shows that the accuracy and the precision of the whole system may be suitable in interventions where the tolerance on the target is around 5 mm (Table 3.1). The time required for the segmentation and registration is around one second. The navigation system adds also the visual component, that is the possibility to show in real time other structures segmented together with the target structures, such as vessels or ducts.

3.5 Comments

There are many methods for the so-called registration problem but, going toward the OR applications, the noose is tightening. In fact, while the extended theoretical works on the subject is very refined and sometimes complex, as we have seen, the systems that are used nowadays in the OR are based on very simple registration methods that were

not developed specially for this purpose. On the other hand, the registration algorithms seen in the previous chapter, cannot handle the registration in the OR.

As we have seen in the current chapter, the landmark based solution gives good results and has a very low complexity. The approach based on landmarks can be extended to handle even dense surface data, but it is very likely, in this case, to lose the correspondences. In this case, the only known technique is the ICP algorithm. In effect, most of the solutions on the market employ this technique. In our opinion this solution can not achieve the desired results as it is an iterative method and the local minima are still an issue.

We tried to use an alternative method, the principal-axes registration, that does not run the same risk of the local minima, and that can be portable to the OR. The intention was to go in the direction of finding appropriate algorithms for designing a navigation system in the OR, based on registration. Toward an in-vivo test, we have validated the method on realistic phantoms with shape similar to the human organs and radiological properties that mimic the human tissue. The algorithms for processing the radiological data were oriented toward the real-time performance. The large amount of data induced us to use parallelized computation for the segmentation process and to ignore the radiologic information in the registration process.

The registration algorithm shall handle not only the rigid component, such as the principal axes algorithm but, after the global rigid alignment, it should handle also the local deformations. The current state of the art algorithms use mathematical models to handle deformations such as thin plate spline or other radial basis functions, therefore they should be adapted and validated for the real-time intra-operative application.

Even if the principal-axes method can work well in an in-vitro set-up, its implementation in an in-vivo set-up would require further refinements.

Chapter 4

Generalized Shape Spaces

We can only see a short distance ahead, but we can see plenty there that needs to be done.

Alan Turing

The studies we have performed in the chapter 2 and in the chapter 3 showed that the automatic solution, such as the three-steps algorithms, is not feasible in the OR, meanwhile the solution based on landmarks and/or processing of the data requires the manual interaction.

We have seen that the second solution is the preferred one, but it is still troublesome especially when the correspondences are not known. In fact, when there are only a few landmarks, the solution may be not as accurate as we expect and, increasing the number of landmarks, as for instance in the case of binary images, the correspondence between landmarks is lost.

In medical image registration, if we consider the images as functions of intensities (equation (2.1)), there is no information about the geometric structures we want to register. On the other hand, if we extract the geometric structures by means of segmentation algorithms (equation (3.1)) we loose information about the intensity or about other structures that were not identified by the segmentation process. On both approaches we have to make a compromise. Considering the fact that the datasets we are dealing with are very large and a registration algorithm suitable for the operating room must be executed in real-time we exclude the use of the whole radiologic

Parts of this chapter are based upon: B. Maris, P. Fiorini. *Procrustes method without correspondences for registration of unstructured data*, 2014, (submitted).

information of the dataset as input to the registration algorithm. The extraction of geometric features is still a research argument but, at least for the pre-operative data, we have the possibility of good segmentation since the dataset is available beforehand and an automatic segmentation may be validated by means of the intervention of an expert person.

What is left after the pre-processing step will be a set of points that represents similar structures in two or more images and our task is to find correspondences between pair of points.

The ICP algorithm is the only well accepted technique to recover correspondences and transformation but is limited by the presence of many local minimum solutions.

In this chapter we try to see how much we have to simplify the problem in order to have a valid solution for the operating room. Once we define the problem, we give a robust solution in both theoretical and practical form.

Another contribution that we present in this chapter is the introduction of the *generalized shape space* framework, which is a method that extend the Procrustes analysis to cases when the homologies are not known *a priori*.

4.1 Correspondences, registration and the Procrustes method

Each set of points that we want to register is called *shape* and the registration task is to find a way to align two or more shapes. The Procrustes method of shape comparison arose as a way of superimposing point-sets with known correspondence. A modern and complete study of the Procrustean metric and shape manifolds was presented by Kendall in [57] and was further extended by the author in [58].

A variety of objects can be represented as point sets in \mathbb{R}^d , where d is usually 2 or 3. One is often presented with the problem of deciding whether two of these point sets, and/or the corresponding underlying objects or manifolds, represent the same geometric structure or not. In the case of correspondence, we are interested in the transformation that relates one form to another. A connected fundamental question is: what conditions must a set of points verify in order to faithfully represent an object? Another question is: what kind of similarity we want to achieve between the objects, i.e. rigid or non rigid? The easiest hypothesis is when the correspondences are known, there is a small amount of noise in the point cloud representation and the transformation is rigid. The closed form solution was given by Shonemann in [99].

The mathematical aspects of the theory of shapes are most of the times not practical in the implementation of algorithms for object recognition and matching. Point-based methods to register surfaces which brings relatively dense point sets into correspondence have become popular, as we have seen, with the introduction of the iterative closest point (ICP) method [10], [130].

In the context of the determination of correspondences and transformation in successive steps several authors built extensions or generalizations of this approach.

While retaining the basic iterative principle of ICP, Rangarajan et. al in [93] formulated a variant of the Procrustes distance between two discrete sets of points in which the correspondence maps are unknown a priori. Their algorithm alternates between calculating optimal rotations and determining correspondence maps. For every fixed rotation \mathbf{R} , it computes the association matrix M between the two sets of points $A = \{A_i\}$ and $B = \{B_j\}$, minimizing the average of the square residuals $\sum_{i,j} M_{ij} \|\mathbf{R}A_i - B_j\|^2$, under the soft constraint that M is indeed a measure of coupling, i.e the values of M elements are between 0 and 1, each row sums to 1, and the 1 value of an element M_{ij} means the perfect correspondence between A_i and B_j . As is the case of the original ICP, this algorithm can also converge to a local rather than a global minimum, and the correspondence maps can still be discontinuous and distorting. Memoli [75] provides theoretical exposition of a similar functional in the context of Gromov-Hausdorff distances between shapes. Ghosh et. al [36] used a similar framework with a smooth surface deformation mechanism together with closest-point maps to determine both correspondence maps and the transformations in an alternating iterative procedure. Their algorithm requires user initialization which may influence the outcome; the way correspondences are assigned can lead to a deformation mechanism that produces a distorting and/or discontinuous map between the two surfaces.

Shapiro and Brady [105] match feature points on the basis of consistent same-space distances by an eigen-analysis technique, following the inter-image distance-based matching technique of Scott and Longuet-Higgins [100]. The solution presented in [100] has a very elegant implementation founded on a well-conditioned eigen-vector solution which involved no iteration, but does not handle large rotations and may become unstable for some value of the parameters. Conversely, Shapiro et al. in [105] introduce a modal shape description to handle also the rotations and the instability, but their solution lacks of the formal proof and, in our tests, does not always provide the results we would have expected.

Boyer et al. [14] introduced the concept of continuous Procrustes distance and

proved that it provides a true metric for two-dimensional surfaces derived from anatomical structures embedded in a three-dimensional space.

Jian and Vemuri in [53] reformulate the task of point set registration as the problem of aligning two Gaussian mixture models (GMM) such that a statistical discrepancy measure between the two corresponding mixtures is minimized. Another probabilistic approach that uses GMM and a closed form solution to establish the correspondences using the expectation maximization algorithm was given by Myronenko and Song in [82].

It is worth to mention the paper of Belongie et al. [9] that introduces the *shape context* concept used to measure the similarity between shapes in two steps: 1) solving for correspondences, 2) using the correspondences to estimate an aligning transform.

These are only a few of the works dealing with the transformation-correspondence problem and, from the first to the last citation, they are all facing the same 'egg-chicken' dilemma: find the correspondences to get the transformation or find the transformation to get the correspondences. The best solutions are of course those where the transformation allowed is non-rigid. In this case we may find perfect correspondences using a mathematical transformation, but one question arises: is it not true that if we allow for non-rigid correspondence we may align any two given objects, such as an apple and a pear or a mouse and an elephant?

It is important to consider the applications and the input-output of the algorithms dealing with the correspondence-registration combination. The input of such algorithms may represent different pose of the same object or views of an object subject to a set of changes such as deformations. The data describing the input can be images or geometric features such as points or triangle meshes. The correspondence could be global in the case of an image dataset or local in the case of geometric features matching. The output of such algorithms gives the transformation that register the datasets on one hand and the correspondence or the measure of match on the other hand.

In the OR applications the rigid transformation is a fundamental output even if we are comparing deformable objects because the intervention requires the presence of other objects whose position is known with respect to the intra-operative imaging system through a calibration process.

If the input is given as a set of geometric features to be matched with another set of features, for instance in the point-set registration, once non-rigidity is allowed, there is an infinite number of ways to map one set onto another. The smoothness constraint is necessary because it discourages mapping that are too arbitrary. As we saw in chapter

3 one of the simplest and most used measures is the space integral of the square of the second order derivatives of the mapping function (see equation (3.7)). This leads to the thin-plate-spline (TPS) function. Introduced by Bookstein in [11] for surface registration in medical imaging and morphometry, and formally described by Wahba in [121] these functions are currently used by most non-rigid registration algorithms of point sets (see also [9], [24], [53], [82]).

The TPS function is easy to compute and implement and it has the advantage to decouple the transformation into an affine part and a non-linear deformable part. In situations where there is no shearing and scaling we can constrain the affine transformation to a rigid one.

Usually, all the non-rigid algorithms first find a common reference system of the two datasets, then proceed with the deformation of one dataset in order to fit the second dataset. Our main concern is that, once we introduce a non-rigid deformation, even if the initial alignment is not satisfactory, the algorithm will yield a very good alignment because of the freedom of the deformation. By keeping trace of the rigid alignment, as in the TPS case, we may assert the goodness of the alignment and the usefulness of the registration. For instance, Chui and Rangarajan in [24], implemented the registration with a deterministic annealing scheme to optimize the correspondence matrix by updating the transformation parameters. The algorithm is clearly attempting to solve the matching problem using a coarse to fine approach. Global structures such as the center of mass and principal axis are first matched, followed by the non-rigid matching of local structures. This means that the rigid alignment will be given by the alignment of the geometric moments of the two data sets considered. This solution, already encountered in chapter 3 has a number of drawbacks such as the sensibility to outliers, noise, occlusions but also to the deformation which is our main concern.

Lipman and Funkhouser in [66] used a different approach for the computation of correspondences of approximately and/or partially isometric surfaces. They employed the Mobius transformation and random sampling to vote for the best correspondences for each triplets of points extracted from each of the two datasets to be registered. In their approach, the datasets given as 3D meshes, which are genus-0 surfaces, were conformally mapped to a sphere.

All of these approaches work well on synthetic models but our preoccupation is how useful are they in practical applications. For instance, in medical image applications, where there are multiple acquisitions of the same anatomical area, if we use

different sensors (e.g. computed tomography CT and ultrasound US) how can an algorithm distinguish between noise and deformation, or how can a deformable algorithm take into account the outliers? Since the deformation is modeled by a mathematical model, such as TPS, we need to be sure of the correct rigid alignment, which is, in most of the cases, not guaranteed.

There is also a concern about the computational cost of these algorithm and the numerous parameters that must be solved/known in advance.

In the following, we propose a simplification of the hypothesis of the problem in order to have a solution that is completely controllable by the theoretical development.

The simplified hypothesis is when we have two sets of points represented in different coordinate systems without knowing the homologies; the theoretical part consider the ideal case when no noise is present, while in the practical implementation we study how robust is this assumption in the presence of noise.

4.2 Shape-spaces and shape-manifolds: intro

The departing point of our generalized shape space method is the the work of Kendall [57] who introduces the theory of shapes. The theory of shapes is concerned about of k labelled points $\mathbf{x}_1, \dots, \mathbf{x}_k$ or $k - ad$ in an Euclidean space \mathbb{R}^d , where $k \geq 2$. Normally, the centroid of the k points will serve as the origin, and the scale will be such that the sum of the squared distance of the points from that origin will be equal to unity. Informally, the *shape* is 'what is left when the differences which can be attributed to translations, rotations and dilatations have been quotiented out'.

By ignoring the translation, scaling and rotation, it has been proven (see [58]) that the shape space denoted by the symbol Σ_d^k has the dimension:

$$\mathbf{d}_d^k = d(k - 1) - \frac{1}{2}d(d - 1) - 1. \quad (4.1)$$

This is because of the constraints on the total number of degrees of freedom (DOF) which is reduced accordingly by the DOF of the translations d , rotations $\frac{1}{2}d(d - 1)$ and scaling. Equation (4.1) holds provided that $k \geq d + 1$.

The author introduces a norm and a metric topology on the shape space deriving the *shape-manifolds*. The distance between two $k - ad$ on the shape manifold Σ_d^k is called the *procrustes distance*.

Starting from the description of the shape by using a set of points, a natural extension of this theory is to quotient out also the effect of the labeling of the points. In

our approach we keep the significance of the scaling. Accordingly, the definition of the shape or generalized shape becomes: 'a generalized shape is what is left when the differences which can be attributed to translations, rotations and permutations have been quotiented out'.

In this chapter we study the implications of such an assumption, we build the theoretical basis and we give some practical results. The invariance of a point set under the action of rotations, translations and permutations will be studied in accordance with the set of distances between each possible pair of points in the set.

4.3 Theoretical foundations

Let us fix a coordinate system in \mathbb{R}^d .

Definition 1. If we denote by $O(d)$ the group of the $d \times d$ orthogonal matrices, and by S_k the group of all permutations of $\{1, \dots, k\}$, a set of k points $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subset \mathbb{R}^d$ is equivalent to $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\} \subset \mathbb{R}^d$, in the sense of **generalized shape**, if and only if there exists $R \in O(d)$, a vector $\mathbf{t} \in \mathbb{R}^d$ and a permutation $\pi \in S_k$ such that:

$$R\mathbf{x}_i + \mathbf{t} = \mathbf{y}_{\pi(i)} \forall i = 1, \dots, k. \quad (4.2)$$

Notice that the introduced notion is an equivalence relation.

Definition 2. We write $[X]$ the equivalence class of X and we call it **generalized shape**.

For any two vectors $\mathbf{x} = (x^1, \dots, x^d)$ and $\mathbf{y} = (y^1, \dots, y^d)$ in \mathbb{R}^d we denote by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^d x^i y^i$ their scalar product and by $\|\mathbf{x} - \mathbf{y}\| = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle^{1/2}$ the Euclidean distance between them. The Euclidean distance defines a metric space called the Euclidean space.

Definition 3. For any set $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ of vectors in \mathbb{R}^d , the **center of mass** is given by the vector $\bar{\mathbf{x}} = \frac{1}{k}(\mathbf{x}_1 + \dots + \mathbf{x}_k)$. The set $\bar{X} = \{\mathbf{x}_1 - \bar{\mathbf{x}}, \dots, \mathbf{x}_k - \bar{\mathbf{x}}\}$ is called the **centered coordinates**.

In the following we shall use the notation X to represent a set of k points $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^d$ describing a generalized shape and, in the same time, a $d \times k$ matrix with columns $\mathbf{x}_1, \dots, \mathbf{x}_k$.

If we sum over i and divide by k (4.2) we have:

$$R \frac{1}{k} \sum_{i=1}^k \mathbf{x}_i + \mathbf{t} = \frac{1}{k} \sum_{i=1}^k \mathbf{y}_i,$$

therefore we can express \mathbf{t} as:

$$\mathbf{t} = \bar{\mathbf{y}} - R\bar{\mathbf{x}}, \quad (4.3)$$

so we can rewrite (4.2) as:

$$R\bar{X} = \bar{Y}_\pi, \quad (4.4)$$

where $\bar{Y}_\pi = \{\mathbf{y}_{\pi(1)} - \bar{\mathbf{y}}, \dots, \mathbf{y}_{\pi(k)} - \bar{\mathbf{y}}\}$.

This formula shows that the only important transformations in the generalized shape definition are the orthogonal transformation, or the rotation, R and the permutation π .

Definition 4. Given $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subset \mathbb{R}^d$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\} \subset \mathbb{R}^d$, we define the *distance* between them as:

$$d(X, Y)_p := \min_{\pi \in S_k} \|d(\mathbf{x}_i, \mathbf{x}_j) - d(\mathbf{y}_{\pi(i)}, \mathbf{y}_{\pi(j)})\|_p, \quad (4.5)$$

that is the minimum of the L_p norm over all the permutations of elements in Y .

This distance ranges from the L_1 norm

$$d(X, Y)_\infty := \min_{\pi \in S_k} \sum_{1 \leq i, j \leq k} |d(\mathbf{x}_i, \mathbf{x}_j) - d(\mathbf{y}_{\pi(i)}, \mathbf{y}_{\pi(j)})| \quad (4.6)$$

to L_∞ norm

$$d(X, Y)_1 := \min_{\pi \in S_k} \max_{1 \leq i, j \leq k} |d(\mathbf{x}_i, \mathbf{x}_j) - d(\mathbf{y}_{\pi(i)}, \mathbf{y}_{\pi(j)})| \quad (4.7)$$

Observe that, since S_k is finite, it makes sense to define (4.5) as a minimum rather than as an infimum.

We use this definition when the correspondences between the points are not known and the computation of the registration transformation is not required.

In the case the correspondences are known but the registration transformation between shapes is not known, the distance between them is given by the following definition:

De nition 5. If $\bar{\mathbf{x}} = \frac{1}{k} \sum_{i=1}^k \mathbf{x}_i$ and $\bar{\mathbf{y}} = \frac{1}{k} \sum_{i=1}^k \mathbf{y}_i$ are the centroids of the point-sets X and Y , and the centroid sizes are given by $S_X = [\frac{1}{k} \sum_{i=1}^k (\mathbf{x}_i - \bar{\mathbf{x}})^2]^{\frac{1}{2}}$ and $S_Y = [\frac{1}{k} \sum_{i=1}^k (\mathbf{y}_i - \bar{\mathbf{y}})^2]^{\frac{1}{2}}$, respectively, the **classical Procrustes distance** $d_P(X, Y)$ between X and Y is defined by

$$d_P(X, Y) = \inf_{\mathbf{M} \in SE(d)} \left(\sum_{i=1}^k \left\| \frac{\mathbf{x}_i \mathbf{M}}{S_X} - \frac{\mathbf{y}_i}{S_Y} \right\|^2 \right)^{\frac{1}{2}}, \quad (4.8)$$

where $SE(d)$ is the special euclidean group of translation and rotations in \mathbb{R}^d .

If we embed the shapes in the Euclidean space \mathbb{R}^d we can imagine the shape as the orbit of all isometric subspaces with the given shape.

De nition 6. We say that two shapes X and Y , embedded in the Euclidean space \mathbb{R}^d , are **isometric** when there exists a bijective mapping $\Phi : X \rightarrow Y$ such that $d(\mathbf{x}_i, \mathbf{x}_j) = d(\Phi(\mathbf{y}_i), \Phi(\mathbf{y}_j))$. Such Φ is an **isometry** between X and Y .

We are interested on whether this isometry exists and how to find it. As the generalized shapes are given as an equivalence class modulo rotations, translation and permutations, the isometry, in the case exists, is given by the composition of the three functions. Since it is clear that rotations and translations are always isometries we are wandering in which case the permutations still lead to isometry. To do this we will work directly with the distance distributions between pairwise points in each set.

De nition 7. Given a set of points $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^d$, we call **distance distribution matrix** the $k \times k$ matrix whose entries are given by the pairwise distances $D_{i,j}^X = \|\mathbf{x}_i - \mathbf{x}_j\|$

A reduced form of the distance distribution matrix is the distance distribution vector.

De nition 8. Given a set of points $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{R}^d$, we call **distance distribution vector** the $\mathbb{R}^{\binom{k}{2}}$ vector whose entries are given by the pairwise distances:

$$V^X = (V_{1,2}, \dots, V_{1,k}, V_{2,1}, \dots, V_{k-1,k}) \quad (4.9)$$

where $V_{i,j}^X = \|\mathbf{x}_i - \mathbf{x}_j\|$ with $1 \leq i < j \leq k$

De nition 9. $\{V_{1,2}, \dots, V_{1,k}, V_{2,1}, \dots, V_{k-1,k}\}$ is called **distance distribution set**.

Note that the distance distributions are invariant under rigid motions. The permutation of k indices will yield a permutation of the distance distributions. The next sections

will analyze when the distance distributions suffice to characterize the orbit of a generalized shape and how we can recover the rigid transformation and the correspondences between two generalized shapes.

Definition 10. Given a set of points $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subset \mathbb{R}^d$, we call **Gram matrix** the $k \times k$ matrix whose entries are given by the inner products $G_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$. The Gram matrix may also be given as the matrix product $X^T X$.

Definition 11. A $n \times n$ matrix M is called **symmetric** if $M_{ij} = M_{ji}$ for all $i, j = 1, \dots, n$. A $n \times n$ matrix M is called **positive semidefinite** if for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^T M \mathbf{x} \geq 0$,

The Gram matrix is positive semidefinite and symmetric, and every positive semidefinite matrix is the Gram matrix for some set of vectors. Further, in finite-dimensions it determines the vectors up to isomorphism, i.e. any two sets of vectors with the same Gram matrix must be related by a single unitary matrix as we shall state in the next lemma.

Lemma 1. For any two X and Y $d \times k$ matrices, if their Gram matrices are equal, i.e. $X^T X = Y^T Y$, then there is a matrix $A \in O(d)$ such that $AX = Y$.

Proof. $X^T X$ is positive semidefinite therefore can be written as $Q \Lambda Q^T$ for $Q \in O(k)$ and a non-negative diagonal matrix Λ . Using the singular value decomposition of X we can write $X = U_X \Sigma Q^T$, where $U_X \in O(d)$ and $\Sigma^T \Sigma = \Lambda$.

Considering $X^T X = Y^T Y$ we may write also $Y = U_Y \Sigma Q^T$, where $U_Y \in O(d)$ and $\Sigma^T \Sigma = \Lambda$.

Then, we can write $AX = Y$ for orthogonal $A = U_Y U_X^T$. □

The next lemma gives useful hints to connect the distance distribution with the Gramian derived from two sets.

Lemma 2. Given the sets $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$, the following statements are equivalent:

- (i) $\|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{y}_i - \mathbf{y}_j\|, \forall 1 \leq i, j \leq k$.
- (ii) $\forall n = 1, \dots, k$ $\langle \mathbf{x}_i - \mathbf{x}_n, \mathbf{x}_j - \mathbf{x}_n \rangle = \langle \mathbf{y}_i - \mathbf{y}_n, \mathbf{y}_j - \mathbf{y}_n \rangle, \forall 1 \leq i, j \leq k$.
- (iii) $\langle \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{x}_j - \bar{\mathbf{x}} \rangle = \langle \mathbf{y}_i - \bar{\mathbf{y}}, \mathbf{y}_j - \bar{\mathbf{y}} \rangle, \forall 1 \leq i, j \leq k$.

Proof. For $\mathbf{p}, \mathbf{q} \in \mathbb{R}^d$ we have:

$$\|\mathbf{p} - \mathbf{q}\|^2 = \langle \mathbf{p} - \mathbf{q}, \mathbf{p} - \mathbf{q} \rangle = \langle \mathbf{p}, \mathbf{p} \rangle + \langle \mathbf{q}, \mathbf{q} \rangle - 2 \langle \mathbf{p}, \mathbf{q} \rangle = \quad (4.10)$$

$$= \|\mathbf{p}\|^2 + \|\mathbf{q}\|^2 - 2 \langle \mathbf{p}, \mathbf{q} \rangle \quad (4.11)$$

For $\mathbf{p} = \mathbf{x}_i - \mathbf{x}_n$ and $\mathbf{q} = \mathbf{y}_i - \mathbf{y}_n$ we derive:

$$\begin{aligned} \|\mathbf{x}_i - \mathbf{x}_j\|^2 &= \|\mathbf{x}_i - \mathbf{x}_n\|^2 + \|\mathbf{x}_j - \mathbf{x}_n\|^2 - 2\langle \mathbf{x}_i - \mathbf{x}_n, \mathbf{x}_j - \mathbf{x}_n \rangle \|\mathbf{y}_i - \mathbf{y}_j\|^2 = \\ &= \|\mathbf{y}_i - \mathbf{y}_n\|^2 + \|\mathbf{y}_j - \mathbf{y}_n\|^2 - 2\langle \mathbf{y}_i - \mathbf{y}_n, \mathbf{y}_j - \mathbf{y}_n \rangle \end{aligned}$$

The last equation and (i) implies (ii). If we have (ii) and we choose $i = j$ in the last equation we obtain $\forall i, n \|\mathbf{x}_i - \mathbf{x}_n\| = \|\mathbf{y}_i - \mathbf{y}_n\|$, therefore (i).

If (ii) holds $\forall i$ we sum and average over i and obtain:

$$\langle \bar{\mathbf{x}} - \mathbf{x}_n, \mathbf{x}_j - \mathbf{x}_n \rangle = \langle \bar{\mathbf{y}} - \mathbf{y}_n, \mathbf{y}_j - \mathbf{y}_n \rangle \forall 1 \leq j \leq k$$

Averaging over j this equation we obtain:

$$\|\bar{\mathbf{x}} - \mathbf{x}_n\| = \|\bar{\mathbf{y}} - \mathbf{y}_n\| \quad (4.12)$$

For $\mathbf{p} = \mathbf{x}_i - \bar{\mathbf{x}}$ and $\mathbf{q} = \mathbf{y}_i - \bar{\mathbf{y}}$ we derive:

$$\begin{aligned} \|\mathbf{x}_i - \mathbf{x}_j\|^2 &= \|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 + \|\mathbf{x}_j - \bar{\mathbf{x}}\|^2 - 2\langle \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{x}_j - \bar{\mathbf{x}} \rangle \|\mathbf{y}_i - \mathbf{y}_j\|^2 = \\ &= \|\mathbf{y}_i - \bar{\mathbf{y}}\|^2 + \|\mathbf{y}_j - \bar{\mathbf{y}}\|^2 - 2\langle \mathbf{y}_i - \bar{\mathbf{y}}, \mathbf{y}_j - \bar{\mathbf{y}} \rangle \end{aligned}$$

The last equation, the equation (4.12) and the equivalence between (i) and (ii) implies (iii).

If we have (iii), then, for $i = j = n$ we obtain $\|\bar{\mathbf{x}} - \mathbf{x}_n\| = \|\bar{\mathbf{y}} - \mathbf{y}_n\|$.

This result, together with the last equation in the (iii) hypothesis, gives (i), therefore (ii) and the proof is complete. \square

The main question is if the distance distribution matrix characterizes in a unique way the generalized shapes.

In the case of labeled points (i.e. in the sense of Kendall's shapes) the answer is yes and it is illustrated by the next theorem.

Theorem 1. *If $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ are sets of points in \mathbb{R}^d and $\|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{y}_i - \mathbf{y}_j\|$, $\forall 1 \leq i, j \leq k$, then there exist a rigid transformation given by $R \in O(d)$ and $\mathbf{t} \in \mathbb{R}^d$ such that $R\mathbf{x}_i + \mathbf{t} = \mathbf{y}_i$ for all i .*

Proof. From Lemma 2 ((i) \Leftrightarrow (iii)), the Gram matrix of the centered coordinates are the same, i.e. $\bar{X}^T \bar{X} = \bar{Y}^T \bar{Y}$.

It follows from Lemma 1 that $\exists R \in O(d)$ such that $R\bar{X} = \bar{Y}$ therefore we have:

$$R(\mathbf{x}_i - \bar{\mathbf{x}}) = \mathbf{y}_i - \bar{\mathbf{y}}, \forall i \text{ or}$$

$$\mathbf{y}_i = R\mathbf{x}_i + \bar{\mathbf{y}} - R\bar{\mathbf{x}}.$$

Let $\mathbf{t} = \bar{\mathbf{y}} - R\bar{\mathbf{x}}$ and the proof is complete. \square

So, in the case of labeled points, i.e. when the correspondences are known, the distance distribution matrix is an invariant that completely characterizes the equivalence class of a shape. We are interested if the same holds in the case of generalized shapes.

The following examples will show that the answer is no, i.e. even if the distribution of distances is the same, there are different configurations that yield the distributions that are not related by a rigid transformation. Before giving these examples we note that the distance distribution of a point-set $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ may be given also as a monotone increasing sequence $(d_1, d_2, \dots, d_{\binom{k}{2}})$, where $\{d_1, d_2, \dots, d_{\binom{k}{2}}\} = \{\|\mathbf{x}_i - \mathbf{x}_j\|\}$ with $1 \leq i < j \leq k$. This sequence is the same as the distance distribution vector up to a permutation and it will be used when we derive the correspondences of points as we will later see (section 4.6).

Consider the two sets of points $X = \{0, 1, 4, 10, 12, 17\}$ and $Y = \{0, 1, 8, 11, 13, 17\}$ in \mathbb{R} . Their distance distribution sequence is given by the ordered vector $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, 17)$ but it is obvious that the point set does not represent the same shape even if we consider the possibility of reflection.

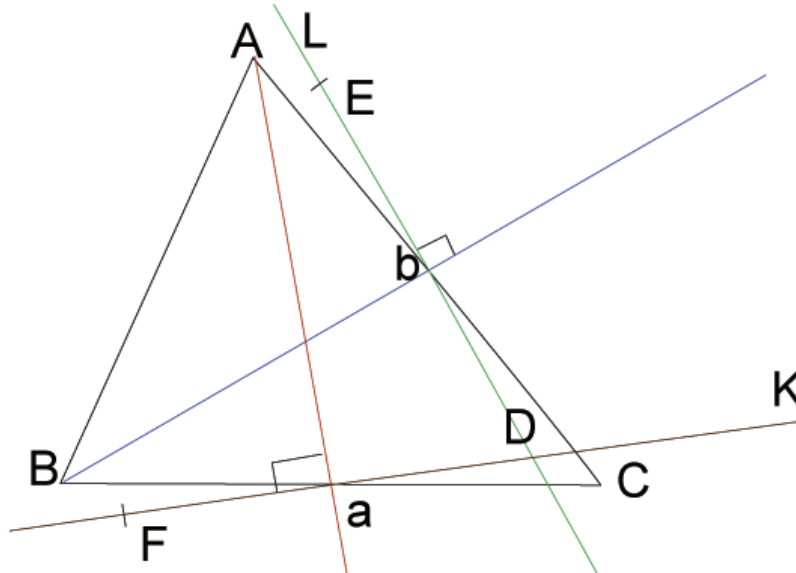


Figure 4.1: The point sets $\{A, B, C, D\}$, $\{A, B, C, E\}$, and $\{A, B, C, F\}$ have the same distribution of distances.

We insert here a more general example in \mathbb{R}^d , $d \geq 2$ taken from [64] and shown in g. 4.1. Given a triangle ABC , let a be the midpoint of BC and b be the midpoint of AC .

Let K be the line through a perpendicular to the line Aa . Let L be the line through b perpendicular to the line Bb . If D is the intersection of L and K , E is the point of L such that $dist(E, b) = dist(D, b)$, and F is the point of K such that $dist(F, a) = dist(D, a)$ then the shapes $\{A, B, C, D\}$, $\{A, B, C, E\}$, and $\{A, B, C, F\}$ have the same distribution distance matrix up to a permutation but they are not isometric. In fact, it is easy to see from the fact that the triangles BDE and ADF are isosceles and $DAEC$ and $DBFC$ are parallelograms that $DA = FA = EC$, $DB = FC = EB$ and $DC = FB = EA$, therefore the 6 distances that form the distance distribution vector of each of the shapes are identical up to a permutation. The point sets given by $\{A, B, C, D\}$, $\{A, B, C, E\}$, and $\{A, B, C, F\}$ do not represent the same generalized shape and we will formally prove in section 4.4.

There are infinite ways to construct shapes with the same distribution of distances as we have seen but having different shapes, therefore we want to know which distance distributions define in a unique way a shape. In the next section we see that most distribution of distances define uniquely a generalized shape.

4.4 Distance distribution and generalized shapes: existence

In this section we analyze some properties about the existence of the generalized shapes using the distribution of distances. Boutin and Kemper [13], [12] characterized the point sets in \mathbf{R}^d that can be determined from their distance distribution. Their results hold in the case when the distribution of pairwise distance are exactly the same. Some of these results will be presented here and they will be integrated in the context of generalized shapes. The section 4.5 will analyze a method to recover the isometric transformation between generalized shapes in the ideal case, when the distribution of distances are exactly the same, and a theoretical study of robustness will be introduced. Finally, the section 4.6 will present a method to compute correspondences between the points of two generalized shapes in the presence of noise.

With the previous result in mind (*Theorem 1*), our approach now is to show which point sets with the same distribution of distances are the same modulo rotations, translations and permutations. More precisely we are interested in the relation between the permutation of the distances as elements of the symmetric group $S_{\binom{k}{2}}$ and the permutations between the point sets.

Denote by C the set of $\binom{k}{2}$ distinct pairs.

$C = \{(i, j) | i \neq j, i, j = 1, \dots, k\}$. This means that $\forall (i, j)$ and $(i', j') \in C$ distinct, the

sets $\{i, j\}$ and $\{i', j'\}$ are also distinct.

Two point sets $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ have the same distance distribution vector if there exists a permutation $\theta \in S_{\binom{k}{2}}$ such that:

$V_{(i,j)}^X = V_{\theta(i,j)}^Y$, $\forall (i, j) \in C$, where $V_{(i,j)}^X = \|\mathbf{x}_i - \mathbf{x}_j\|$ and $V_{(i',j')}^Y = \|\mathbf{y}_{i'} - \mathbf{y}_{j'}\|$ and $\theta(i, j) = (i', j')$ are the components of the distance distribution vector.

Definition 12. θ is a *labeling of the points* if there exists a permutation $\pi \in S_k$ of the indices such that:

$$\theta(i, j) = (\pi(i), \pi(j)), \forall (i, j) \in C. \quad (4.13)$$

The previous definition links the correspondences between pairs of each sets given by two indices $(i, j) \in C$ for X , $\theta(i, j) \in C$ for Y and point correspondences between X and Y given by $\pi(i)$ and $\pi(j)$. The next result shows the immediate correspondence with the generalized shapes.

Corollary 1. Given $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ with the same distribution of distances up to a permutation $\theta \in S_{\binom{k}{2}}$ which is a labeling of points, then $Y \in [X]$ in the sense of generalized shape.

Proof. Consider π from (4.13). From hypothesis $V_{(i,j)}^X = V_{\theta(i,j)}^Y$ but $V_{\theta(i,j)}^Y = V_{(\pi(i), \pi(j))}^Y$ therefore $V_{(i,j)}^X = V_{(\pi(i), \pi(j))}^Y$.

The sets of points X and $Y_\pi = \{\mathbf{y}_{\pi(1)}, \dots, \mathbf{y}_{\pi(k)}\}$ have the same distance distribution vector, therefore by *Theorem 1* there exists $R \in O(d)$ and $\mathbf{t} \in \mathbb{R}^d$ such that $R\mathbf{x}_i + \mathbf{t} = \mathbf{y}_{\pi(i)}$ for all i . This implies $Y \in [X]$. \square

This connection between the permutation of $\binom{k}{2}$ distances and the permutation of k points gives the generalized shape equivalence class. We need to know which $S_{\binom{k}{2}}$ permutations are good permutations in the sense of equation (4.13).

The next theorem, which is fundamental for our argument, prove that a permutation over C is a labeling if it preserves adjacency.

Theorem 2. ([13]) For $k \neq 4$, $\theta \in S_{\binom{k}{2}}$ is a labeling (i.e. induces equivalent shapes modulo rotations, translations and permutations) if and only if $\forall i, j, l$ pairwise distinct indices we have

$$\theta(i, j) \cap \theta(i, l) \neq \emptyset \quad (4.14)$$

Proof. The 'only if' part of the statement is clear from (4.13), so we will show the 'if' direction.

For $k = 3$ the set of 3 distances may yield a unique triangle, therefore every $\theta \in S_{\binom{3}{2}}$ is a good permutation. Same for $k = 2$.

Assume $k \geq 5$ and $\theta \in S_{\binom{k}{2}}$ satisfies (4.14).

Suppose, by contradiction, for pairwise distinct $i, j, l, m \in \{1, \dots, k\}$:

$$\theta(i, j) \cap \theta(i, l) \cap \theta(i, m) = \emptyset.$$

Then, by (4.14) we can write $\theta(i, j) = (a, b)$, $\theta(i, l) = (a, c)$ and $\theta(i, m) = (b, c)$, with $a, b, c \in \{1, \dots, k\}$ pairwise distinct. Now, since we have more than 4 points, we can choose n distinct from i, j, l, m , $\theta(i, n)$ must intersect (a, b) , (a, c) and (b, c) . Since $\theta(i, n)$ only has two elements, it must be one of the sets (a, b) , (a, c) , (b, c) , contradicting the injectivity of θ . Therefore, $\theta(i, j) \cap \theta(i, l) \cap \theta(i, m) \neq \emptyset$.

Then if we fix an i and choose any distinct j, l , $\theta(i, j) \cap \theta(i, l)$ must contain a distinct element a and the above shows that a belongs to any $\theta(i, m)$, where m is distinct from i, j, l . Therefore $\bigcap_{m \neq i} \theta(i, m) = a$ and we can define the map $\pi : \{1, \dots, k\} \rightarrow \{1, \dots, k\}$ such that $\pi(i) = a$. To show that $\pi \in S_n$ we simply need to show that it is injective. To do this, let M_i be the set of all pairs containing i . Then $\theta(M_i) \subseteq M_{\pi(i)}$. But θ is a bijection and $|M_i| = |M_{\pi(i)}|$ therefore $\theta(M_i) = M_{\pi(i)}$.

Let $i \neq j$ with $\pi(i) = \pi(j)$. Then, $M_{\pi(i)} = M_{\pi(j)}$ so $\theta(M_i) = \theta(M_j)$. But θ is a permutation so $M_i = M_j$ therefore $i = j$ so π is injective.

Now consider $\theta(i, j)$. By the above, it contains both $\pi(i)$ and $\pi(j)$, so $\theta(i, j) = (\pi(i), \pi(j))$ that is (4.13). \square

For $k = 4$, *Theorem 2* is not true. In fact, observe that the relation $\theta(i, j) \cap \theta(i, l) \cap \theta(i, m) = \emptyset$ cannot be contradicted because we cannot choose n distinct from i, j, l, m . The theorem becomes true in all cases if we impose the additional condition that, for each pairwise distinct $i, j, l, m \in \{1, \dots, k\}$, $\theta(i, j) \cap \theta(i, l) \cap \theta(i, m) \neq \emptyset$.

A simple counterexample for the $k = 4$ case is given by the point-sets from the previous section 4.3, figure 4.1, where the sets of distances $\{AB, BC, AC, AD, BD, CD\}$ and $\{AB, BC, AC, CE, BE, AE\}$ are the same and the condition (ii) from *Theorem 2* holds, but the point sets are not isometric.

In the case of the example $X = \{0, 1, 4, 10, 12, 17\}$ and $Y = \{0, 1, 8, 11, 13, 17\}$ in \mathbb{R} , $k = 6$, the distributions of the ordered vectors of distances of X and Y are the same: $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, 17)$ so we can check *Theorem 2*. If we denote the points of each set with numbers from 1 to 6, following the natural order, we have:

$V_{(2,3)}^X = V_{\theta(3,4)}^Y$ and $V_{(1,3)}^X = V_{\theta(5,6)}^Y$ so, from *Theorem 2* the two sets do not represent the same generalized shape.

Equipped with this result that allows us to identify good permutations, we show next that if two point sets are sufficiently close and have the same distribution of distances, then they represent the same generalized shape. This result is weaker than desired but gives some intuition on why point sets might be determined by their pairwise distance.

Theorem 3. ([12]) *If $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subset \mathbb{R}^d$ is a set of k points there exists a neighborhood $V(X) \in (\mathbb{R}^d)^k$ of $(\mathbf{x}_1, \dots, \mathbf{x}_k)$ such that if $(\mathbf{y}_1, \dots, \mathbf{y}_k) \in V(X)$ is a configuration with the same distribution of distances vector of X , then the two point configurations belong to the orbit of the same generalized shape.*

Proof. Suppose, by contradiction, that exists a sequence $(Y_n)_{n=1}^{\infty}$ converging to X such that none of Y_n can be mapped to X through a rigid transformation and permutation, but exists a sequence of permutations $(\theta_n)_{n=1}^{\infty} \in S_{\binom{k}{2}}$ such that $V_{(i,j)}^X = V_{\theta_n(i,j)}^{Y_n}$. Since the symmetric group $S_{\binom{k}{2}}$ has a finite number of elements, we can choose θ_1 and construct $(Z_l)_{l=1}^{\infty}$ a subsequence of $(Y_n)_{n=1}^{\infty}$ where $\theta_n = \theta_1$.

Taking the limit $l \rightarrow \infty$, we have $V_{\theta_1^{-1}(i,j)}^X = \lim_{l \rightarrow \infty} V_{(i,j)}^{Z_l}$. Since $(Z_l)_{l=1}^{\infty}$ converges to X , we have $V_{\theta_1^{-1}(i,j)}^X = V_{(i,j)}^X$ or:

$$V_{(i,j)}^X = V_{\theta_1(i,j)}^X \quad (4.15)$$

From the construction of Z_l we have:

$$V_{\theta_1(i,j)}^{Z_l} = V_{(i,j)}^X, \forall i, j, l \quad (4.16)$$

(4.15) and (4.16) imply that $V_{(i,j)}^X = V_{(i,j)}^{Z_l} \forall i, j, l$.

By *Theorem 1*, X and Z_l are the same up to a rigid motion, for every l , so we have a contradiction. □

4.5 Distance distribution and generalized shapes: isometric transformation

In this section we are in the hypothesis of the *Theorem 2*, i.e. if there is a correspondence θ between the distance distributions of two sets, then θ is a labeling. We are interested here to compute the rigid part of the transformation without knowing the point

to point correspondence. This solution to the computation of the correspondences will be given in the next section.

We begin this section with some known general results from linear algebra that we shall use.

Recall that if $\pi \in S_n$ is a permutation of $\{1, 2, \dots, n\}$, then the $n \times n$ matrix P_π :

$$P_\pi = (\mathbf{e}_{\pi(1)} \mathbf{e}_{\pi(2)} \dots \mathbf{e}_{\pi(n)}) \quad (4.17)$$

is the *permutation matrix* associated to π , where \mathbf{e}_j denotes a column vector of length n with 1 in the j th position and 0 in every other position.

Lemma 3. *Let P_π be the permutation matrix associated to π . Then the following holds:*

- (i) P_π is orthogonal and $P_\pi^{-1} = P_\pi^T = P_{\pi^{-1}}$
- (ii) $\det P_\pi = \pm 1$
- (iii) *Multiplying a $d \times n$ matrix X on the right by P_π permutes the columns of X by P_π .*

This lemma is well known and its proof is ignored.

The *eigenvalues* of a $n \times n$ matrix A are the solutions λ of the determinant equation $\det(A - \lambda I_n) = 0$ associated to the *characteristic polynomial* $p_A := \det(A - \lambda I_n)$, where I_n is the $n \times n$ identity matrix.

Associated with each eigenvalue λ of A , there will be non-trivial solutions of the equation $(A - \lambda I_n)\mathbf{x} = 0$. These are called *eigenvectors* of A corresponding to the eigenvalue λ . Thus, if λ is an eigenvalue of A , then there will exist a corresponding eigenvector $\mathbf{v} \neq 0$, which is a non-trivial solution of $(A - \lambda I_n)\mathbf{v} = 0$.

If λ is an eigenvalue of A , then the dimension d_g of the linear space $\text{Null}(A - \lambda I_n) := \{\mathbf{x} | A\mathbf{x} - \lambda\mathbf{x} = 0\}$ is called *geometric multiplicity* of λ , while the multiplicity d_m of λ as root of the characteristic polynomial is called *algebraic multiplicity*.

It is well known that the $d_g \leq d_m$.

Lemma 4. *If P_π is the permutation matrix associated to $\pi \in S_n$ and X and Y are two $n \times n$ positive symmetric semidefinite matrices such that $Y = P_\pi X P_\pi^{-1}$ then:*

- (i) X and Y have the same set of eigenvalues
- (ii) *There exist eigenvalue decompositions of $X = U_X D U_X^{-1}$ and $Y = U_Y D U_Y^{-1}$, such that $P_\pi = U_Y U_X^{-1}$*

Proof. $p_Y(\lambda) = \det(Y - \lambda I)$

Since $I = P_\pi P_\pi^{-1}$ we can write:

$$\begin{aligned} p_Y(\lambda) &= \det(P_\pi(X - \lambda I)P_\pi^{-1}) = \det(P_\pi)\det(X - \lambda I)\det(P_\pi^{-1}) \\ &= \det(X - \lambda I) = P_X(\lambda) \end{aligned}$$

which shows that the eigenvalues of X and Y and their algebraic multiplicity are the same.

Let $X = U_X D U_X^T$ be the eigenvalue decomposition of X . X is symmetric, positive semidefinite therefore D consisting of all eigenvalues of X is real and diagonal.

$$Y = P_\pi X P_\pi^{-1} = P_\pi U_X D U_X^T P_\pi^{-1}$$

$$U_Y = P_\pi U_X \text{ is orthonormal } (U_Y^T = U_X^T P_\pi^T = U_X^T P_\pi^{-1} \text{ and thus}$$

$$Y = U_Y D U_Y^T$$

D is the diagonal matrix of all eigenvalues of X , therefore by (i) of all eigenvalues of Y . \square

Observe that not all the eigen decompositions of X and Y lead to P_π since the decompositions are not unique. The previous lemma ensure only the existence.

We are ready to give a strong result about the connection between the permutations of the distance distribution matrix and the generalized shapes.

Theorem 4. *Given $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ two point sets in \mathbb{R}^d , the following statements are equivalent:*

$$(i) Y \in [X]$$

(ii) $\exists P_\pi$ a permutation matrix, such that $D^X = P_\pi^T D^Y P_\pi$. Moreover, $P_\pi = U_Y U_X^T$, where U_Y and U_X are orthogonal matrices from an eigenvalue decomposition of $\bar{Y}^T \bar{Y}$ and $\bar{X}^T \bar{X}$.

(iii) $\exists P_\pi$ a permutation matrix and $R \in O(d)$, such that $R \bar{X} = \bar{Y} P_\pi$. Moreover, $P_\pi = U_Y U_X^T$, where U_Y and U_X are orthogonal matrices from an eigenvalue decomposition of $\bar{Y}^T \bar{Y}$ and $\bar{X}^T \bar{X}$.

Proof. Lemma 2 says that the correspondences between the elements of the distance distribution matrix are the same as the correspondences between the Gram matrix of the centered coordinates ((i) \Leftrightarrow (iii)), therefore if we show that there exists $P_\pi \in S_k$ such that $\bar{X}^T \bar{X} = P_\pi^T \bar{X}^T \bar{X} P_\pi$, the same relationship will hold between D^X and D^Y .

$Y \in [X] \Leftrightarrow$ exists $R \in O(d)$, a vector $\mathbf{t} \in \mathbb{R}^d$ and a permutation $\pi \in S_k$ such that:

$$R \mathbf{x}_i + \mathbf{t} = \mathbf{y}_{\pi(i)} \text{ for all } i = 1, \dots, k.$$

Let P_π be the permutation matrix associated with π .

If we denote $Y_\pi = \{\mathbf{y}_{\pi(1)}, \dots, \mathbf{y}_{\pi(k)}\}$, we have

$$Y_\pi = Y P_\pi \text{ and}$$

$$\bar{Y}_\pi = \bar{Y} P_\pi. \tag{4.18}$$

From hypothesis $\|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{y}_{\pi(i)} - \mathbf{y}_{\pi(j)}\| \forall i, j = 1, \dots, k$, if and only if, by *Lemma 2*, $\langle \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{x}_j - \bar{\mathbf{x}} \rangle = \langle \mathbf{y}_{\pi(i)} - \bar{\mathbf{y}}, \mathbf{y}_{\pi(j)} - \bar{\mathbf{y}} \rangle, \forall 1 \leq i, j \leq k$.

This is equivalent to $\bar{X}^T \bar{X} = \bar{Y}_\pi^T \bar{Y}_\pi$ or, using (4.18) $\bar{X}^T \bar{X} = P_\pi^T \bar{Y}^T \bar{Y} P_\pi$.

The formula of P_π follows from *Lemma 4*.

We can rewrite (ii) as $\bar{X}^T \bar{X} = (\bar{Y} P_\pi)^T \bar{Y} P_\pi$, therefore by *Lemma 1* this is true if and only if there exists an orthogonal matrix R such that $R\bar{X} = \bar{Y} P_\pi$.

We have so far a relation between centered coordinates \bar{X} and \bar{Y} together with the matrix $R \in O(d)$ and $\pi \in S_k$.

The equivalence as generalized shapes between X and Y follows the same reasoning as in *Theorem 1*, putting $\mathbf{t} = \bar{\mathbf{y}} - R\bar{\mathbf{x}}$.

□

Notice again that the eigenvalue decomposition is not unique therefore the previous result ensures only the existence of the matrix P_π and does not help in finding the permutation or the orthogonal matrix that relates the two point sets.

In our assumption, the permutation that realizes the correspondence between the point sets is not known, so we want to find the orthogonal matrix that relates the two point sets without knowing the correspondences. The next theorem will introduce a way to find this matrix. The complete solution of this problem will be given as an algorithmic method.

The $d \times d$ matrix $\bar{Y}\bar{Y}^T$ does not depend on the permutation of the elements of Y , since for all $\pi \in S_k$, $\bar{Y}_\pi \bar{Y}_\pi^T = \bar{Y} P_\pi P_\pi^T \bar{Y}^T = \bar{Y}\bar{Y}^T$. This allows us to establish the next result that will be the base of the algorithm that finds the rigid transformation between two sets representing the same generalized shape without knowing the correspondences between the point sets.

Theorem 5. *Given $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ two point sets in \mathbb{R}^d with $Y \in [X]$, then the matrices $\bar{X}\bar{X}^T$ and $\bar{Y}\bar{Y}^T$ have the same eigenvalues, including their algebraic multiplicities. If the eigenvalues are all distinct we denote by S the $d \times d$ matrix of the \mathbf{s}_i eigenvectors of $\bar{X}\bar{X}^T$ written as columns, T the $d \times d$ matrix of the \mathbf{t}_i eigenvectors of $\bar{Y}\bar{Y}^T$ written as columns and $\pi \in S_k$ the permutation determined from $\bar{Y}_\pi = R\bar{X}$. Then we have:*

$$\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle = \delta_i \langle \mathbf{y}_{\pi(j)} - \bar{\mathbf{y}}, \mathbf{t}_i \rangle, \quad (4.19)$$

for all $j = 1, \dots, k, i = 1, \dots, d$ and $\delta_i = \pm 1$.

Moreover, if $\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle \neq 0$ and we denote $\Delta = \text{Diag}(\delta_1, \dots, \delta_d)$, where

$$\delta_i = \frac{\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle}{\langle \mathbf{y}_{\pi(j)} - \bar{\mathbf{y}}, \mathbf{t}_i \rangle} \quad (4.20)$$

than $R = T\Delta S^T$.

Proof. From *Theorem 4* it follows that if $Y \in [X]$ then $\exists R \in O(d)$ such that $\bar{Y}_\pi = R\bar{X}$.

We have $\bar{Y}\bar{Y}^T = R\bar{X}P_\pi^T P_\pi \bar{X}^T R^T = R(\bar{X}\bar{X}^T)R^T$.

If we decompose the real, symmetric matrix $\bar{X}\bar{X}^T$ using the eigendecomposition $\bar{X}\bar{X}^T = S\Lambda S^T$ then:

$$\bar{Y}\bar{Y}^T = (RS)\Lambda(RS)^T. \quad (4.21)$$

RS is orthogonal as product of orthogonal matrices so (4.21) shows that $\bar{X}\bar{X}^T$ and $\bar{Y}\bar{Y}^T$ have the same eigenvalues, including their algebraic multiplicities and RS is a matrix of eigenvectors for $\bar{Y}\bar{Y}^T$.

Since the eigenvalues are distinct and knowing that the geometric multiplicity is less than the algebraic multiplicity, the dimension of $\text{Null}(\bar{X}\bar{X}^T - \lambda_i I_d)$ is one therefore, for all $i = 1, \dots, d$ if \mathbf{t}_i is an eigenvector of $\bar{Y}\bar{Y}^T$ then $\mathbf{t}_i = \pm R\mathbf{s}_i$.

In the case $\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle \neq 0$ denoting $\delta_i = \frac{\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle}{\langle \mathbf{y}_{\pi(j)} - \bar{\mathbf{y}}, \mathbf{t}_i \rangle}$ and $\Delta = \text{Diag}(\delta_1, \dots, \delta_d)$, we can write $R = T\Delta S^T$.

Since the inner product is invariant to isometries and $\bar{Y}_\pi = R\bar{X}$, with R orthogonal we can write:

$$\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle = \langle R(\mathbf{x}_j - \bar{\mathbf{x}}), R\mathbf{s}_i \rangle = \pm \langle \mathbf{y}_{\pi(j)} - \bar{\mathbf{y}}, \mathbf{t}_i \rangle,$$

and the proof is completed. \square

Remark that if $\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle = 0$ for all $j = 1, \dots, k$, δ_i cannot be determined from (4.20). In this case $\bar{X}^T \mathbf{s}_i = 0$, therefore $\bar{X}\bar{X}^T \mathbf{s}_i = 0$. This means \mathbf{s}_i is the eigenvector corresponding to the eigenvalue 0. This remark allows us to give the following corollary.

Corollary 2. *If the eigenvalues are all distinct and non-zero, then $\exists j = 1, \dots, k$ such that $\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle \neq 0$, and therefore:*

$$\delta_i = \frac{\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle}{\langle \mathbf{y}_{\pi(j)} - \bar{\mathbf{y}}, \mathbf{t}_i \rangle}$$

Following this corollary, the matrix $\Delta = \text{Diag}(\delta_1, \dots, \delta_d)$ is completely determined only when π is known, while we set out to achieve the computation of R without knowing π . The next corollary will solve this problem by giving a way to compute Δ .

Let's first denote, for each $i = 1, \dots, d$, A_i^- the sets:

$$A_{i-}^X = \{|\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle| : \langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle < 0, j = 1, \dots, k\}, i = 1, \dots, d.$$

$$A_{i+}^X = \{|\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle| : \langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle > 0, j = 1, \dots, k\}, i = 1, \dots, d.$$

In a similar way we define A_{i-}^Y and A_{i+}^Y .

Corollary 3. *If $Y \in [X]$ and none of the eigenvalues of the matrix $\bar{X}\bar{X}^T$ is zero, then, for each $i = 1, \dots, d$ one of the two cases happens:*

- (i) $A_{i-}^X = A_{i-}^Y$ and $A_{i+}^X = A_{i+}^Y$ and thus $\delta_i = 1$
- (ii) $A_{i-}^X = A_{i+}^Y$ and $A_{i+}^X = A_{i-}^Y$ and thus $\delta_i = -1$

In this way the matrix $\Delta = \text{Diag}(\delta_1, \dots, \delta_d)$ is determined without knowing the permutation π .

We conclude this section with the following algorithm that, in most cases, solves the rigid transformation between two elements of the same generalized shape, in the case where the point to point correspondences are not known.

Algorithm 1 Orthogonal transformation between two generalized shapes

- 1: Input X and Y two k point sets in \mathbb{R}^d
 - 2: Compute $\bar{X}\bar{X}^T$ and $\bar{Y}\bar{Y}^T$ and their eigenvalues $\lambda_1, \dots, \lambda_d$ and μ_1, \dots, μ_d
 - 3: If the eigenvalues are not all distinct in each set then the algorithm is inconclusive
 - 4: If the set $\{\lambda_1, \dots, \lambda_d\}$ is different from the set $\{\mu_1, \dots, \mu_d\}$ then $Y \notin [X]$
 - 5: Compute \mathbf{s}_i and $\mathbf{t}_i, i = 1, \dots, d$ the sets of eigenvectors of $\bar{X}\bar{X}^T$ and $\bar{Y}\bar{Y}^T$
 - 6: If $\exists i$ such that the sets $\{|\langle \mathbf{x}_j - \bar{\mathbf{x}}, \mathbf{s}_i \rangle| : j = 1, \dots, k\}$ and $\{|\langle \mathbf{y}_j - \bar{\mathbf{y}}, \mathbf{t}_i \rangle| : j = 1, \dots, k\}$ do not coincide then $Y \notin [X]$
 - 7: Compute $A_{i-}^X, A_{i+}^X, A_{i-}^Y, A_{i+}^Y$
 - 8: If $\exists i = 1, \dots, d$ such that $A_{i-}^X \neq A_{i-}^Y \vee A_{i+}^X \neq A_{i+}^Y$ and $A_{i-}^X \neq A_{i+}^Y \vee A_{i+}^X \neq A_{i-}^Y$ then $Y \notin [X]$
 - 9: For all $i = 1, \dots, d$ define $\delta_i = 1$ when $A_{i-}^X = A_{i-}^Y \wedge A_{i+}^X = A_{i+}^Y$ and $\delta_i = -1$ when $A_{i-}^X = A_{i+}^Y \wedge A_{i+}^X = A_{i-}^Y$
 - 10: Build the matrix $R = T\Delta S^T$, where $\Delta = \text{Diag}(\delta_1, \dots, \delta_d)$
-

In general, the eigenvalues of a matrix cannot be computed exactly, as they are roots of a polynomial, this making our algorithm impractical. The next section will introduce, along with a method to find the point to point correspondences, a robust solution of the computation of the transformation matrix between two shapes. However, if the distribution of distances of two point sets is the same up to a threshold, we can compare the eigenvalues of the Gramian matrices of the centered coordinates by using a very small threshold ε . We have done the same to compare the sets $A_{i_-}^X, A_{i_+}^X, A_{i_-}^Y, A_{i_+}^Y$.

Complexity

The Algorithm 1 requires the computation of $\bar{X}\bar{X}^T$ (complexity at most $O(d^2k)$, its eigenvalue decomposition (complexity $O(d^3)$, the computation of δ_i (complexity $O(dk)$), the computation of R (complexity at most $O(d^3)$). Since $d \ll k$ the resulting complexity is at most $O(dk^2)$.

Robustness

The next lemma will show the equivalence between the distribution of distances in the presence of noise and the correspondent values of the elements in the Gram matrix of the centered coordinates.

Proposition 1. *Let $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ two point sets in \mathbb{R}^d . Then the following are equivalent:*

- (i) $\exists \varepsilon > 0$ such that $\forall i, j = 1, \dots, k \left| \|\mathbf{x}_i - \mathbf{x}_j\|^2 - \|\mathbf{y}_i - \mathbf{y}_j\|^2 \right| \leq \varepsilon$.
- (ii) $\exists \delta > 0 \forall i, j = 1, \dots, k \left| \langle \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{x}_j - \bar{\mathbf{x}} \rangle - \langle \mathbf{y}_i - \bar{\mathbf{y}}, \mathbf{y}_j - \bar{\mathbf{y}} \rangle \right| \leq \delta$

Proof. (i) \Rightarrow (ii)

As in Lemma 2 we substitute in (4.10) \mathbf{p} with $\mathbf{x}_i - \mathbf{x}_n$, \mathbf{q} with $\mathbf{x}_j - \mathbf{x}_n$ and the same for $\mathbf{y}_i - \mathbf{y}_n$, $\mathbf{y}_j - \mathbf{y}_n$ to derive:

$$\langle \mathbf{x}_i - \mathbf{x}_n, \mathbf{x}_j - \mathbf{x}_n \rangle = \frac{1}{2} \left(\|\mathbf{x}_i - \mathbf{x}_n\|^2 + \|\mathbf{x}_j - \mathbf{x}_n\|^2 - \|\mathbf{x}_i - \mathbf{x}_j\|^2 \right),$$

$$\langle \mathbf{y}_i - \mathbf{y}_n, \mathbf{y}_j - \mathbf{y}_n \rangle = \frac{1}{2} \left(\|\mathbf{y}_i - \mathbf{y}_n\|^2 + \|\mathbf{y}_j - \mathbf{y}_n\|^2 - \|\mathbf{y}_i - \mathbf{y}_j\|^2 \right),$$

If we average both identities over i , subtract them and use (i) we obtain:

$$\left| \langle \bar{\mathbf{x}} - \mathbf{x}_n, \mathbf{x}_j - \mathbf{x}_n \rangle - \langle \bar{\mathbf{y}} - \mathbf{y}_n, \mathbf{y}_j - \mathbf{y}_n \rangle \right| \leq \frac{3}{2} \varepsilon.$$

Now, we average the last inequality over j and we derive:

$$\left| \|\mathbf{x}_n - \bar{\mathbf{x}}\|^2 - \|\mathbf{y}_n - \bar{\mathbf{y}}\|^2 \right| \leq \frac{3}{2} \varepsilon, n = 1, \dots, k. \quad (4.22)$$

We apply again (4.10) for $\mathbf{p} = \mathbf{x}_i - \bar{\mathbf{x}}$, $\mathbf{q} = \mathbf{x}_j - \bar{\mathbf{x}}$ and for $\mathbf{p} = \mathbf{y}_i - \bar{\mathbf{y}}$, $\mathbf{q} = \mathbf{y}_j - \bar{\mathbf{y}}$:

$$\langle \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{x}_j - \bar{\mathbf{x}} \rangle = \frac{1}{2} \left(\|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 + \|\mathbf{x}_j - \bar{\mathbf{x}}\|^2 - \|\mathbf{x}_i - \mathbf{x}_j\|^2 \right) \quad (4.23)$$

$$\langle \mathbf{y}_i - \bar{\mathbf{y}}, \mathbf{y}_j - \bar{\mathbf{y}} \rangle = \frac{1}{2} \left(\|\mathbf{y}_i - \bar{\mathbf{y}}\|^2 + \|\mathbf{y}_j - \bar{\mathbf{y}}\|^2 - \|\mathbf{y}_i - \mathbf{y}_j\|^2 \right) \quad (4.24)$$

We subtract (4.23) and (4.24) and we put $\delta = \varepsilon/2$ to get:

$$\left| \langle \mathbf{x}_i - \bar{\mathbf{x}}, \mathbf{x}_j - \bar{\mathbf{x}} \rangle - \langle \mathbf{y}_i - \bar{\mathbf{y}}, \mathbf{y}_j - \bar{\mathbf{y}} \rangle \right| \leq \delta$$

(ii) \Rightarrow (i)

We substitute in (ii) j with i we have:

$$\left| \|\mathbf{x}_i - \bar{\mathbf{x}}\|^2 - \|\mathbf{y}_i - \bar{\mathbf{y}}\|^2 \right| \leq \delta$$

If we subtract (4.23) and (4.24) and use this last inequality we obtain, for $\varepsilon = 4\delta$:

$$\left| \|\mathbf{x}_i - \mathbf{x}_j\|^2 - \|\mathbf{y}_i - \mathbf{y}_j\|^2 \right| \leq \varepsilon$$

□

4.6 Distance distribution and generalized shapes: correspondences and registration

4.6.1 Distance distribution permutation

So far, we have seen how to verify the existence of the generalized shape when we know the distance distribution matrix or the distance distribution vector and how to find the rigid transformation between two elements of the same generalized shape, without knowing the correspondences.

Algorithm 1 shows how to find the registration matrix when we know the point sets and the fact that they belong to the same generalized shape. This result is not guaranteed in the presence of a high level of noise making it not usable for most practical applications.

In this section we give a method to compute the correspondences between two point sets belonging to the same generalized shape in the presence of noise and a method to derive the rigid transformation that relates the two point sets using the correspondences.

The problem to solve is, given $X = \{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_k\}$ two point sets in \mathbb{R}^d with $Y \in [X]$, to find the permutation $\pi \in S_k$ such that:

$$RX + \mathbf{t} = YP_\pi, \quad (4.25)$$

where $R \in O(d)$, $\mathbf{t} \in \mathbb{R}^d$ and P_π is the permutation matrix associated to π .

Since $O(k!)$ different correspondence permutations are possible amongst the two point sets, brute force search is intractable.

As before, the invariant we chose, to find the right correspondences, is given by the distance distribution vector as defined in section 4.3.

Theorem 2 proved in what cases the distance distribution vector defines a generalized shape in a unique way.

We shall derive the permutation π from the permutation $\gamma \in S_{\binom{k}{2}}$ that 'align' the two distribution distance vectors $V^X = (V_{1,2}^X, \dots, V_{k-1,k}^X)$ and $V^Y = (V_{1,2}^Y, \dots, V_{k-1,k}^Y)$.

$$V^X P_\gamma = V^Y. \quad (4.26)$$

Equation (4.26) holds in the ideal case, when no noise is present. In the practical application we may relax the exact correspondence and search for a permutation matrix P_γ of dimension $k(k-1)/2$ such that:

$$P_\gamma = \arg \min_{P_\phi} \|V^X P_\phi - V^Y\|_1 \quad (4.27)$$

The use of the Manhattan distance or L_1 norm in (4.27) is coherent with the definition of the distance between generalized shapes given in the section 4.3 by the equation 4.6. Through the rest of this chapter, when there are no ambiguities, we shall use the notation $\|\cdot\|_1 = |\cdot|$.

The L_1 norm works better in the presence of noise and is equivalent with L_2 norm, in the sense of equivalence between norms. Considering $\|\cdot\|_2 \leq \|\cdot\|_1$, the equation (4.27) gives an upper bound also for the L_2 norm.

Since there are $(k(k-1)/2)!$ ways to arrange a shape vector, we need an efficient mode to solve (4.27). We shall further see that this process takes $O(N \log N)$, where $N = k(k-1)/2$, by ordering each of the shape vectors.

Lemma 5. *If $a_1 \leq a_2 \leq \dots \leq a_N$, $b_1 \leq b_2 \leq \dots \leq b_N$ and $\pi : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$ is a permutation of the indices, then we have:*

$$\begin{aligned} |a_1 - b_1| + \dots + |a_N - b_N| &\leq \\ |a_1 - b_{\pi(1)}| + \dots + |a_N - b_{\pi(N)}|, \end{aligned} \quad (4.28)$$

Proof. We use a proof by induction on N .

The proposition $P(N)$ is given by (4.28).

• $P(2)$:

$$|a_1 - b_1| + |a_2 - b_2| \leq |a_1 - b_2| + |a_2 - b_1|, \quad (4.29)$$

with $a_1 \leq a_2$ and $b_1 \leq b_2$.

Suppose $b_1 \leq a_1$, the other case is symmetric.

Then (4.29) becomes

$$\begin{aligned} a_1 - b_1 + |a_2 - b_2| &\leq |a_1 - b_2| + a_2 - b_1 \\ \Leftrightarrow |a_2 - b_2| &\leq |a_2 - a_1| + |a_1 - b_2| \end{aligned}$$

that is the triangle inequality.

• $P(N - 1) \Rightarrow P(N)$

If $\pi(1) = 1$, the first and the last term in (4.28) are identical so we are in the $P(N - 1)$ case and the proof is completed.

If $\pi(1) = i, i \in \{2, \dots, N\} \implies \exists j \in \{2, \dots, N\}$ such that $\pi(j) = 1$.

We are in the hypothesis of $P(2)$: $a_1 \leq a_j (a_j \neq 1)$ and $b_1 \leq b_j (b_j \neq 1)$

$$\implies |a_1 - b_1| + |a_j - b_i| \leq |a_1 - b_i| + |a_j - b_1|$$

From the right side of (4.28):

$$\begin{aligned} &|a_1 - b_{\pi(1)}| + \dots + |a_N - b_{\pi(N)}| \\ &= |a_1 - b_i| + \dots + |a_j - b_1| + \dots + |a_N - b_{\pi(N)}| \\ &\geq |a_1 - b_1| + \dots + |a_j - b_i| + \dots + |a_N - b_{\pi(N)}| \end{aligned}$$

We define the function $\pi' : \{2, \dots, k\} \mapsto \{2, \dots, N\}$ as $\pi'(l) = \pi(l)$, for $l \neq j$ and $\pi'(j) = i$.

It is easy to see that π' is a bijective mapping, therefore a permutation of $N - 1$ elements. From $P(N - 1)$, it follows:

$$\begin{aligned} &\implies |a_1 - b_1| + |a_2 - b_2| + \dots + |a_N - b_N| \\ &\leq |a_1 - b_1| + |a_2 - b_{\pi'(2)}| + \dots + |a_j - b_i| + \dots + |a_N - b_{\pi'(N)}| \\ &= |a_1 - b_1| + \dots + |a_j - b_i| + \dots + |a_N - b_{\pi(N)}| \quad (\text{qed}). \end{aligned}$$

□

Corollary 4. For any π_A and π_B permutation of N indices and $a_1 \leq a_2 \leq \dots \leq a_N$, $b_1 \leq b_2 \leq \dots \leq b_N$ it holds:

$$\sum_{i=1}^N |a_i - b_i| \leq \sum_{i=1}^N |a_{\pi_A(i)} - b_{\pi_B(i)}| \quad (4.30)$$

Theorem 6. The solution of (4.27) is given by:

$$P_\gamma = P_{\pi_X} P_{\pi_Y}^T, \quad (4.31)$$

where P_{π_Y}, P_{π_X} are the permutations matrices that order the vectors V^Y and V^X respectively.

Proof. From the hypothesis that vectors $V^X P_{\pi_X}$ and $V^Y P_{\pi_Y}$ are ordered, it follows the Corollary 4 holds:

$$|V^X P_{\pi_X} - V^Y P_{\pi_Y}| \leq |V^X P'_{\pi_X} - V^Y P'_{\pi_Y}|$$

for every and all permutations P'_{π_X}, P'_{π_Y} .

Multiplying on the right with $P_{\pi_Y}^T$ the expression $V^X P_{\pi_X} - V^Y P_{\pi_Y}$ we have:

$$|V^X P_{\pi_X} - V^Y P_{\pi_Y}| = |V^X P_{\pi_X} P_{\pi_Y}^T - V^Y|$$

from which the conclusion follows. \square

We are now able to find the minimum distance between two shape vectors that represent in an unique mode two point-sets given in two different coordinate systems regardless of the point ordering.

Once we have the correspondence between the shape vectors the next step is to find the correspondences between the points. This will be described in an algorithmic form in the next section.

4.6.2 Determining correspondences

Point-set correspondences in a closed form solution

The correspondences between the indices of the point-sets X and Y , when $Y \in [X]$ will be given again as a permutation matrix. To do this we have to associate the two indices i, j that give the distance $D_{i,j}^X = \|\mathbf{x}_i - \mathbf{x}_j\|$ in the distance distribution matrix (or alternatively $D_{i,j}^Y$) to a unique index l in the distance distribution vector and vice versa, to each of the indexes of the distance distribution vector two indexes in the point-set.

In this second case we have two solutions considering the order of the two vectors that compose the distance.

From equation (4.9), considering that i represents the index of a row and j represents the offset inside the i -th row in the distance distribution matrix, the mapping $(i, j) \mapsto l$ is given by:

$$(i, j) \mapsto (i-1)(2k-i)/2 + j - i \quad (4.32)$$

Conversely, if we denote the distance distribution vector index l , the problem is to find the correspondent (i, j) couple of indexes, where again i represents the row and j represents the off-set in the distance distribution matrix.

In this case, i is given by the smallest positive number that satisfies:

$$l \leq ik - i(i+1)/2 \quad (4.33)$$

Transforming (4.33) in a second order inequality:

$$i^2 - (2k-1)i + 2l \leq 0 \quad (4.34)$$

we observe that the sum and the product of the roots are positive, therefore the solution for the i -th index is given by the smallest integer bigger than the smallest root of the associated second order equation:

$$i = \left\lceil k - 1/2 - \sqrt{(2k-1)^2 - 8l/2} \right\rceil \quad (4.35)$$

From (4.35) and (4.32) it follows:

$$j = l + i - (i-1)(2k-1)/2 \quad (4.36)$$

We are now able to write the algorithm that finds the correspondences of two point-sets having the same cardinality, regardless of the coordinate system where each of the sets are represented and of their ordering.

Algorithm 2 Correspondences Algorithm

- 1: Input X, Y two k point sets in \mathbb{R}^d , $Y \in [X]$
 - 2: Compute the distance distribution vectors V^X, V^Y
 - 3: Compute the permutations π_Y, π_X that orders V^X, V^Y
 - 4: $P_Y \leftarrow P_{\pi_X} P_{\pi_Y}^T$
 - 5: **For** $i_X := 1 \text{ to } k - 1$
 - 6: (4.32) gives the indexes l_X^1, l_X^2 in V^X corresponding to $(i_X, i_X + 1), (i_X, i_X + 2)$
 - 7: Compute $l_Y^1 = P_Y(l_X^1)$ and $l_Y^2 = P_Y(l_X^2)$
 - 8: Use (4.35) and (4.36) and find (i_Y^1, j_Y^1) and (i_Y^2, j_Y^2) corresponding to l_Y^1, l_Y^2
 - 9: The correspondence of X_i is Y_{i_c} ,
where $i_c = \{i_Y^1, j_Y^1\} \cap \{i_Y^2, j_Y^2\}$
 - 10: **End For**
 - 11: Find the last correspondence
-

Observe that the line 9 of algorithm 2 is guaranteed by the *Theorem 2*.

Point-set correspondences with noise

The algorithm 2 gives the closed form solution of the correspondence of the two point sets X and Y representing the same generalized shape.

This solution holds in the ideal case, when no noise is present. In the case of noisy data there is no guarantee to find a common index i_c in the row 9-th of the algorithm 2. To solve this we use a technique of voting, associating to each couple $(\mathbf{x}_i, \mathbf{y}_j)$ an increasing vote for each possible correspondence, therefore we build an association matrix. To find the correspondence matrix M we simply extract the maximum value in each row of the association matrix.

The detailed algorithm is as following:

Algorithm 3 Correspondence Algorithm with noise

- 1: Input X, Y two k point sets in \mathbb{R}^d , $Y \in [X]$
 - 2: Initialize a $k \times k$ matrix M to zero
 - 3: Compute the vectors V^X, V^Y
 - 4: Compute the permutations π_Y, π_X that orders V^X, V^Y
 - 5: $P_\gamma \leftarrow P_{\pi_X} P_{\pi_Y}^T$
 - 6: **For** $l_A := 1 \text{ to } k(k-1)/2$
 - 7: Compute (i_X, j_X) corresponding to l_X and (i_Y, j_Y) corresponding to $\gamma(l_X)$
 - 8: Increase of 1 $M(i_X, i_Y), M(i_X, j_Y), M(j_X, i_Y), M(j_X, j_Y)$
 - 9: **End For**
 - 10: Set the largest element of each row of M to 1 and the others to 0
-

4.6.3 Correspondence test and results

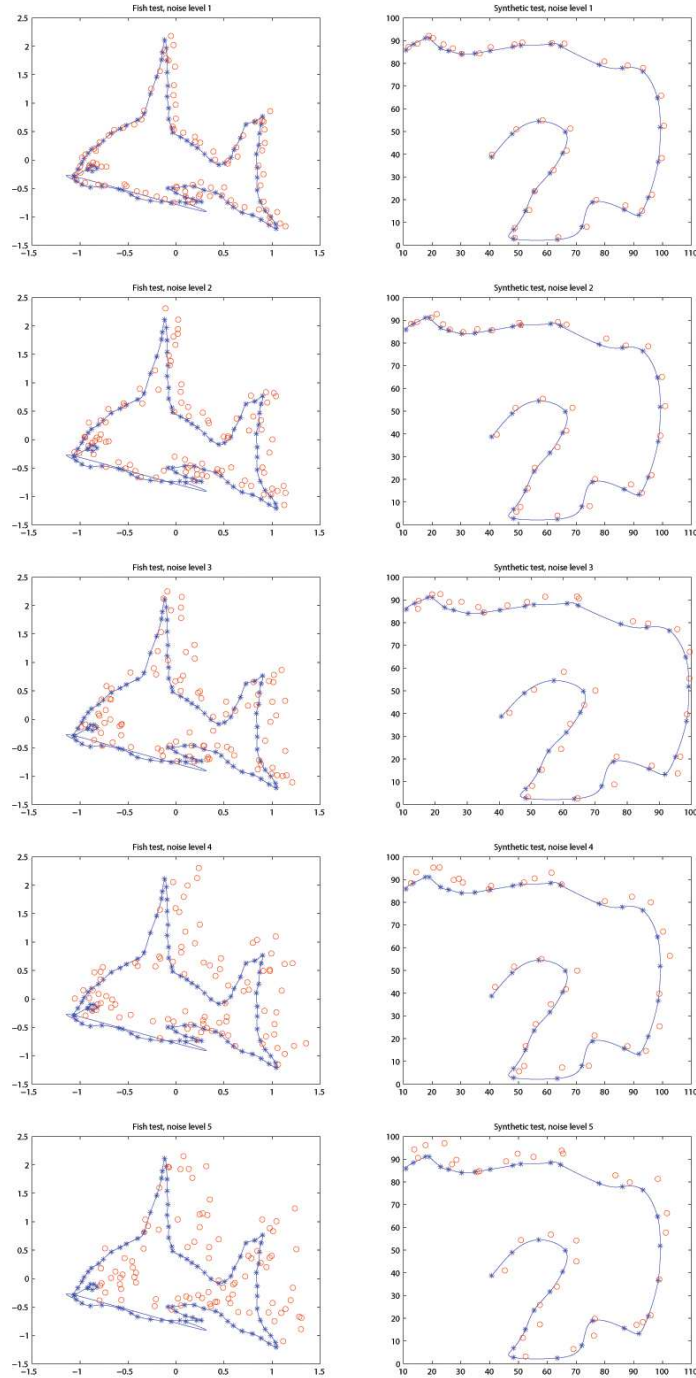


Figure 4.2: Testing 2D data for empirical robustness evaluation. On each of the images the points marked with blue '*' represent the model (joined by a continuous line obtained through interpolation) meanwhile the points marked with red 'o' represent the model transformed by a random permutation and different levels of noise. The fish model on the left column follows [53], [82].

In this section we present the results of the correspondence algorithm. We run tests with different levels of noise and we compute, for each of the tests, the percentage of good correspondences.

The point sets used for testing are depicted in the figures 4.2 and 4.3.

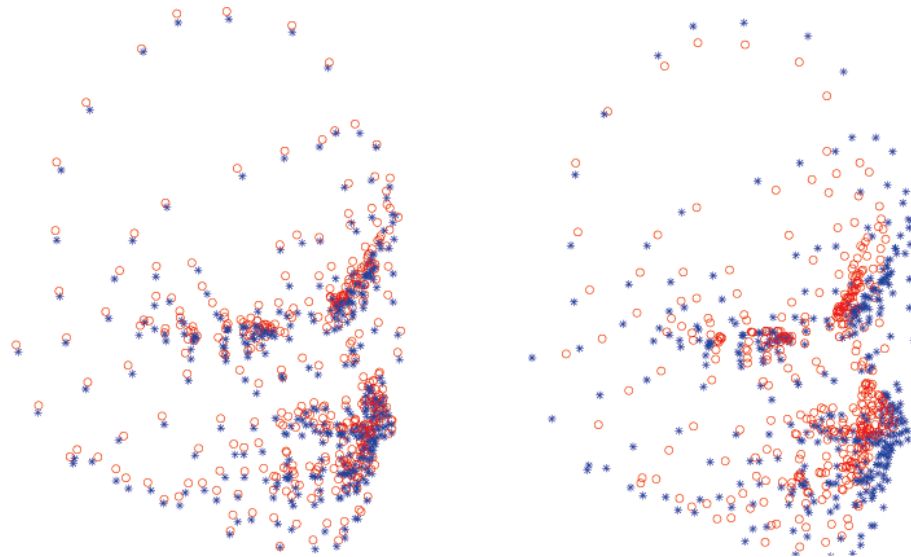


Figure 4.3: Two 3D datasets with different noise levels. The points marked with blue '*' represent the model and the points marked with red 'o' represents the model transformed by a random permutation and different levels of noise.

Figure 4.2 shows the original point-set aligned along a continuous line obtained by interpolation, while the 'o' points are obtained by adding random noise to a permutation of the original dataset. Figure 4.3 depicts the 3D model dataset denoted by the '*' blue points. The dataset was permuted and random noise was added (red circles) before the correspondence algorithm was launched.

In both 2D and 3D tests, since we compute the correspondences relying only on the shape vector, which is invariant to isometries, the permutations of the noisy dataset suffice to test the robustness of our algorithm.

Even if the tests done are very simple and synthetic, we reiterate here that we are addressing applications where there is noise and small deformations without a clear distinction between the two (considered that data usually represent the surface or the border of an object in the real world and is acquired as unstructured point cloud). Our main interest in recovering correspondences is to get the best rigid alignment between the data sets.

The noise we added for the purposes of this work is composed by pseudorandom

Test no-points no	Noise level	Good corresp.
1-34	1%	88%
1-34	2%	76%
1-34	3%	65%
1-34	4%	56%
1-34	5%	41%
2-38	1%	89%
2-38	2%	70%
2-38	3%	62%
2-38	4%	53%
2-38	5%	44%
sh-91	1%	72%
sh-91	2%	65%
sh-91	3%	58%
sh-91	4%	37%
sh-91	5%	34%
face 3D-392	1%	65%
face 3D-392	2%	48%
face 3D-392	3%	37%
face 3D-392	4%	35%
face 3D-392	5%	24%

Table 4.1: Correspondences test

values drawn from the standard normal distribution with mean 1 and different standard deviations (see Table 4.1). The noise level is given by the standard deviation $\sigma_i, i \in \{1, \dots, 5\}$ and represents a percentage fraction of the maximum extension of the point set X , that is $\sigma_i = i \max_{\mathbf{x}_i, \mathbf{x}_j \in X} \|\mathbf{x}_i - \mathbf{x}_j\| / 100$.

The computation of the rigid transformation in the case of noisy data will be addressed in the next section. In this section we compute the correspondences between two point sets X and $Y = XP_\pi + N(\mu, \sigma^2)$, where P_π is a random permutation of $|X|$ elements, N is the normal distribution of mean $\mu = 0$ and variance σ^2 (see figures 4.2 and 4.3).

As we can see from the results (Table 4.1), the number of wrong correspondences increases with the amount of noise. This is an obvious observation but we are more concerned whether we are able to recover the right registration transformation from the correspondences we found. The answer is yes as we see in the next section and, in this case, we can update the correspondences after the registration using the nearest neighbor technique. Another observation is that good correspondences are strongly related to the highest value in the correspondences matrix M from the algorithm 3.

Keeping in mind that we only need 3 correspondences to find the parameters for the rigid alignment we may keep only a part of the good correspondences to find the best rigid alignment. The selection order of this correspondence will be given by the highest value in the correspondences matrix.

4.6.4 Registration tests and results

Recovering the rigid transformation from the correspondences and the distance between the shapes

The main reason of our work is not only to find the correspondences but to find the best rigid alignment between two point-sets $X = \{x_1, \dots, x_k\}, Y = \{y_1, \dots, y_k\}, X, Y \subset \mathbb{R}^d$ that describe a shape in different coordinate systems. In the ideal case the correspondences are guaranteed under the form of a permutation matrix Π (section 4.6.2) and the translation vector $\mathbf{t} \in \mathbb{R}^d$ and the $d \times d$ rotation matrix \mathbf{R} are completely determined in 2D or 3D by a closed form solution using the SVD (singular value) decomposition [2].

For the completeness of presentation we briefly give here the solution in our context, for a detailed explanation see also [55].

The problem is to find \mathbf{R} and \mathbf{t} such that:

$$\Pi Y = \mathbf{R}X + \mathbf{t}, \quad (4.37)$$

In (4.37) and the following X and Y are given as $d \times k$ matrices.

If we use the definition of the Procrustes distance (definition 5), we may reformulate (4.37) as a minimization problem:

$$d_P(X, Y) = \min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^k \|\Pi_i Y - \mathbf{R}x_i - \mathbf{t}\|^2, \quad (4.38)$$

where Π_i is the i -th row of the matrix Π .

Using the centered coordinates (definition 3, section 4.3), the translation vector is given by (4.3) and ((4.37) becomes:

$$\Pi \bar{Y} = \mathbf{R} \bar{X}, \quad (4.39)$$

We denote by $\|A\|_F = \sqrt{\text{trace}(A^T A)} = \sqrt{\sum_{i,j} a_{i,j}^2}$ the Frobenius norm and we obtain an instance of the original Procrustes problem:

$$d_P(X, Y) = \min_{\mathbf{R}} \|\Pi\bar{Y} - \mathbf{R}\bar{B}\|_F^2 \quad (4.40)$$

The solution of this instance is given in closed form by:

$$\mathbf{R} = V \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \det(VU^T) \end{pmatrix} U^T \quad (4.41)$$

where $\Pi\bar{Y}\bar{X}^T = UDV^T$ is the SVD of the matrix $\Pi\bar{Y}\bar{X}^T$.

When no scaling is present, this solution holds for point-sets of cardinality $k \geq 3$.

In the second case, when noise is present, there is no guarantee, as shown in section 4.6.3, that all the correspondences are good. We need a method to filter out the bad correspondences and to recover the registration transformation from the good correspondences.

Even if the percentage of good correspondences lowers with the increase of the noise level, the minimum number of the good correspondences we need to find the rigid transformation parameters is only 3. To accelerate the computation, when we know *a priori* about the presence of noise, we may use the results obtained in section 4.6.3 and keep only the best percentage of correspondences. For each correspondences of 3 points we compute the transformation parameters. Since we do not have all the correspondences, the use of the Procrustes distance is not feasible, therefore we have chosen to compute the *Hausdorff distance* between the two registered shapes:

$$d_H(X, Y') = \max \left\{ \sup_{x \in X} \inf_{y \in Y'} d(x, y), \sup_{y \in Y'} \inf_{x \in X} d(x, y) \right\}, \quad (4.42)$$

where Y' is the point set ΠY after the application of the rigid transformation (figure 4.4).

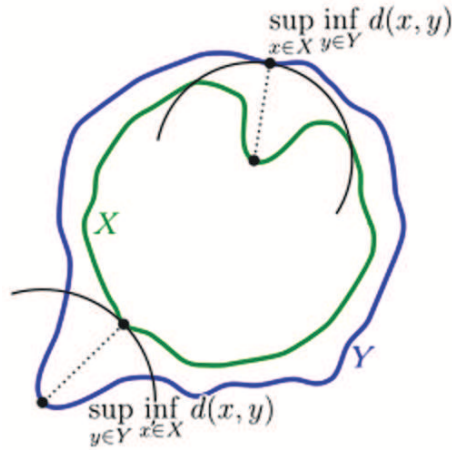


Figure 4.4: The Hausdorff distance

In case no outliers are present (as in our hypothesis), the Hausdorff distance is a good measure of the distance between shapes, while is rarely used in the opposite case: difference in a single sample can make d_H arbitrarily large. In this case an approximation of the Hausdorff distance is given by the ICP distance [10]:

$$d(X, Y') = \sum_{x \in X} d^2(x, Y'), \quad (4.43)$$

where $d^2(x, Y') = \max_{y \in Y'} \|x - y\|_2^2 = \|x - y^*\|_2^2$.

The 3 point correspondences that minimize the Hausdorff distance yield the transformation we are looking for. The computation time is polynomial $O(K)$, where $K = p * k$ and $p \in (0, 1)$ is the percentage of the correspondences we are considering.

The best value for p considering the percentage of good correspondences we have obtained in Table 4.1 is between 0.2 and 0.3. When the noise level is low we may decrease this value because of the high percentage of good correspondences, while when the level of noise is higher we must keep at least 30% of the best correspondences we found, that means $p = 0.3$.

Comparative tests

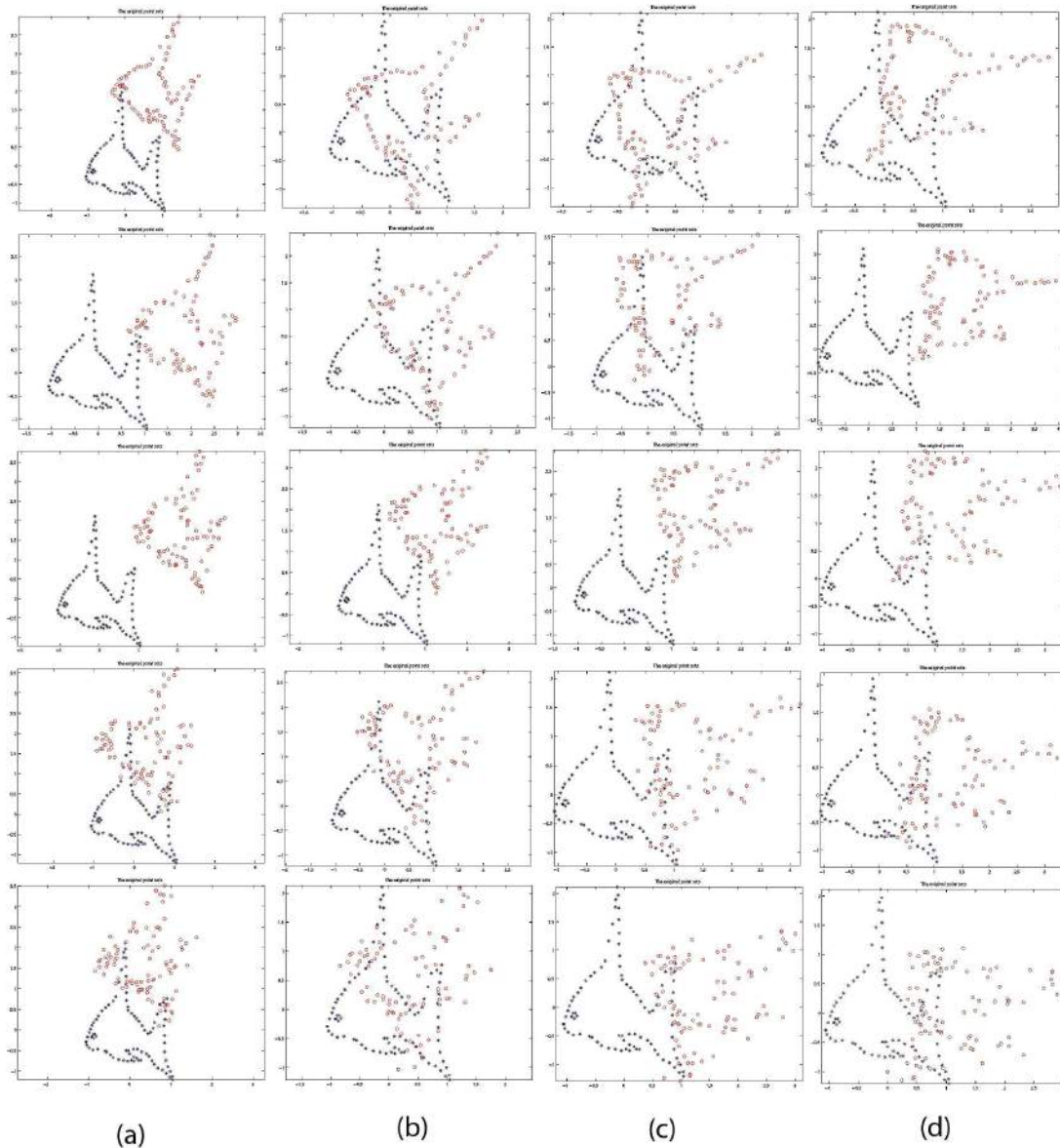


Figure 4.5: The datasets used in the comparison between different algorithms for the registration. (a) The silhouette rotated by 20 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%). (b) The silhouette rotated by 40 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%). (c) The silhouette rotated by 60 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%). (d) The silhouette rotated by 80 degrees and translated by a random vector. The level of noise increases from top (1%) to bottom (5%)

We evaluated the performance of our algorithm in comparison with [53] and [82], that we have already cited before. These algorithms are able to recover the correspondences and a non-rigid transformation.

The fixed and moving point sets used as input to the registration algorithms are represented in figure 4.5. In this set of images the fixed shape is given by the blue stars, while the red circles represent the moving point-set. The scope of the tests is to align the moving point-set with the fixed point set and to evaluate the goodness of the alignment. On each row of the figure 4.5 a different level of noise is added to the moving image. The columns of this image represent the rotation angles between the fixed and the moving image ranging from 20 to 80 degrees with steps of 20 degrees. The translation vector was chosen randomly. It can be seen that, while in the first row the shapes are very similar and distinguishable, in the last row the comparability of the two shapes is difficult even for the human eye. The points of one of the datasets undergone a random permutation.

The results of the registration/correspondence tests are presented in figures 4.6 to 4.10 at the end of this section.

Figure 4.6 gives the results of our algorithm at 3 different noise level. On each of the images the two shapes are registered and the correspondent points are connected by a segment. The Hausdorff distance is given by the length of the segment between the black squares.

Figures 4.7 and 4.8 present the result of the registration/correspondences tests using the coherent point drift (CPD) algorithm ([82]). As before, we give the results of the algorithm at 3 different noise level but, in this case, under each image representing the registered data we have put the unregistered pair with the correspondences to illustrate the situation when the algorithm yields good registration but the correspondences are completely wrong.

In the last figure (figure 4.10) the results of the point set registration using mixture of Gaussians of Jian and Vemuri (JM) algorithm ([53]) are similar to the CPD algorithm, therefore we chose to give only the correspondences of the unregistered images.

The numerical results of the tests we have performed are given in table 4.2. At each level of noise and at each rotation angle we have computed the Hausdorff distance between the registered shapes.

It is easy to see that the accuracy of our algorithm depends on the noise level, while the other algorithms give better results when the level of noise is low and the rigid transformation involves a rotation angle up to 40 degrees but their results are unusable

when the rotation angles goes toward 60 degrees and the noise level increases.

Since our registration algorithm uses only the inter-point distances on each of the datasets, the results are independent of the rotation angle. For low noise level value the registration results are very good, while increasing the noise level there is still a good alignment but the Hausdorff distance increases much already at 3% of noise. This is due to the method for computing the rigid registration from the point correspondences which is dependent on the noise level. We preferred to have a low number of corresponding points to get the rigid transformation in order to maintain low the complexity of the algorithm but increasing the number of corresponding points the registration can be refined.

By moving the points coherently, the CPD algorithm does not modify too much the shape of the figure when applying the deformation, but is unable to yield good results when the rotation angle is 60 degrees or greater (figures 4.7, 4.8 and 4.9 and table 4.2)). We have tested the rigid variant of this algorithm and it performs very well at 60 degrees on all the noise levels ((d) column in the figures 4.7, 4.8 and 4.9, see also the 5-th column of the table 4.2) but it fails to register in all the cases when the rotation angle is 80 degrees ((f) column in the figure 4.7 and figure 4.8 and table 4.2).

On the other hand, the JM algorithm is inclined to deform completely the figure when the rotation angle is bigger than 40 degrees. The distance between the registered shapes is always in line with the level of noise (see table 4.2), but the registration reduces to morphing in all the cases when the rotation angle is greater than 40 degrees (figure 4.10). In these cases one shape is transformed to match the other shape but the point to point correspondences do not reflect the original shape. Figure 4.10 shows this behavior using the line connections between the corresponding registered points.

Considering our application target, the OR, where the position, therefore the rotation, of the shapes to be registered is arbitrary, the CPD and JM algorithms cannot ensure the goodness of the results.

Hausdorff distance					
Noise level	Rot. angle	Our algorithm	Non rigid CPD	Rigid CPD	Non rigid JM
1	20	0.1400	0.0683		0.0711
	40	0.1132	0.0683		0.0768
	60	0.1093	1.5139	0.0678	0.1844
	80	0.1292	1.3939	1.5405	0.1902
2	20	0.2430	0.1465		0.2212
	40	0.2032	0.1464		0.1969
	60	0.2183	1.4244	0.13241	0.2131
	80	0.2430	1.0996	1.5346	0.2524
3	20	0.2935	0.2222		0.2352
	40	0.2829	0.1705		0.4680
	60	0.5427	1.3518	0.1954	0.2575
	80	0.4001	1.3727	1.5145	0.1880
4	20	0.6086	0.2934		0.3112
	40	0.4228	0.3387		0.3232
	60	0.4283	1.2371	0.2724	0.2855
	80	0.5509	1.2439	1.4601	0.2891
5	20	0.7021	0.3157		0.4692
	40	0.6809	0.3184		0.4681
	60	0.6420	1.4163	0.3257	0.3982
	80	0.7417	1.1655	1.3645	0.2176

Table 4.2: Comparison of Error Values after the registration using our algorithm and the other two approaches.

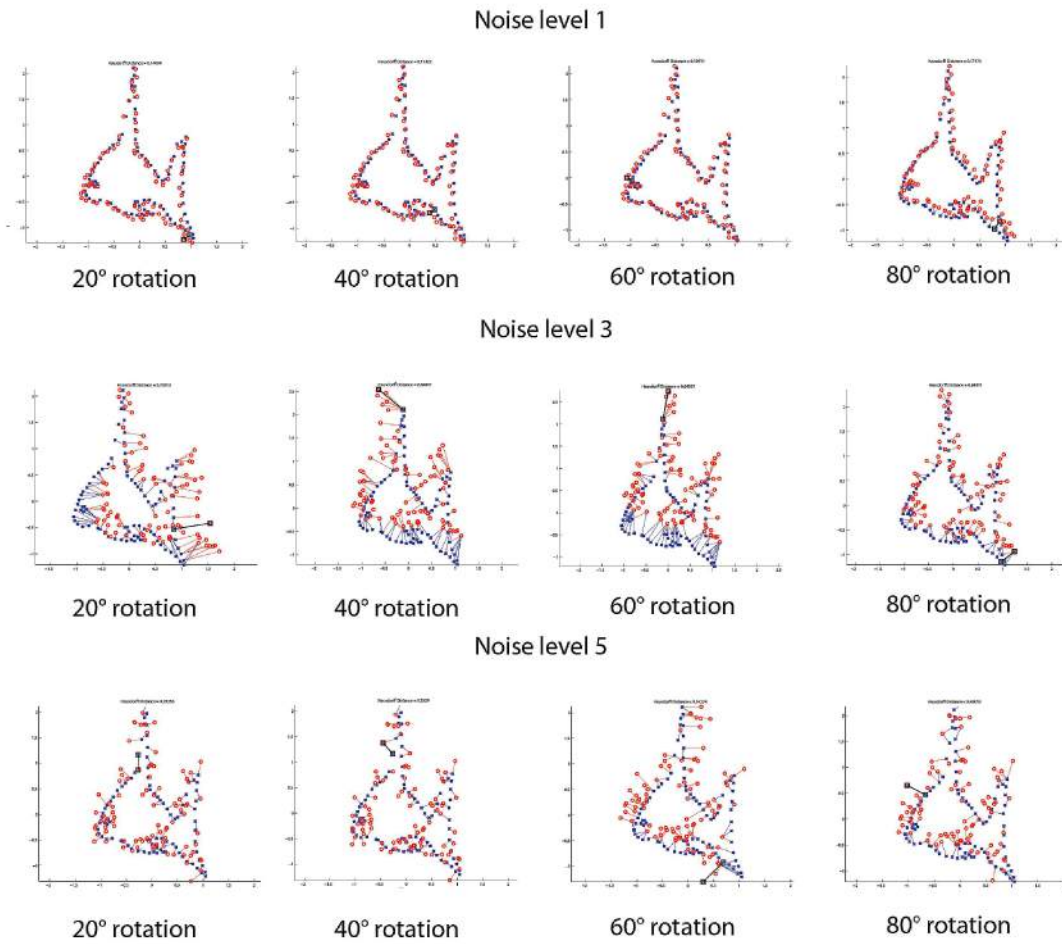


Figure 4.6: The registration using our registration algorithm. On each of the figure the blue stars are the fixed points, meanwhile the red circles represent the moving point-set after the registration took place. The Hausdorff distance is represented by the line between the black squares and its value is given in the table 4.2

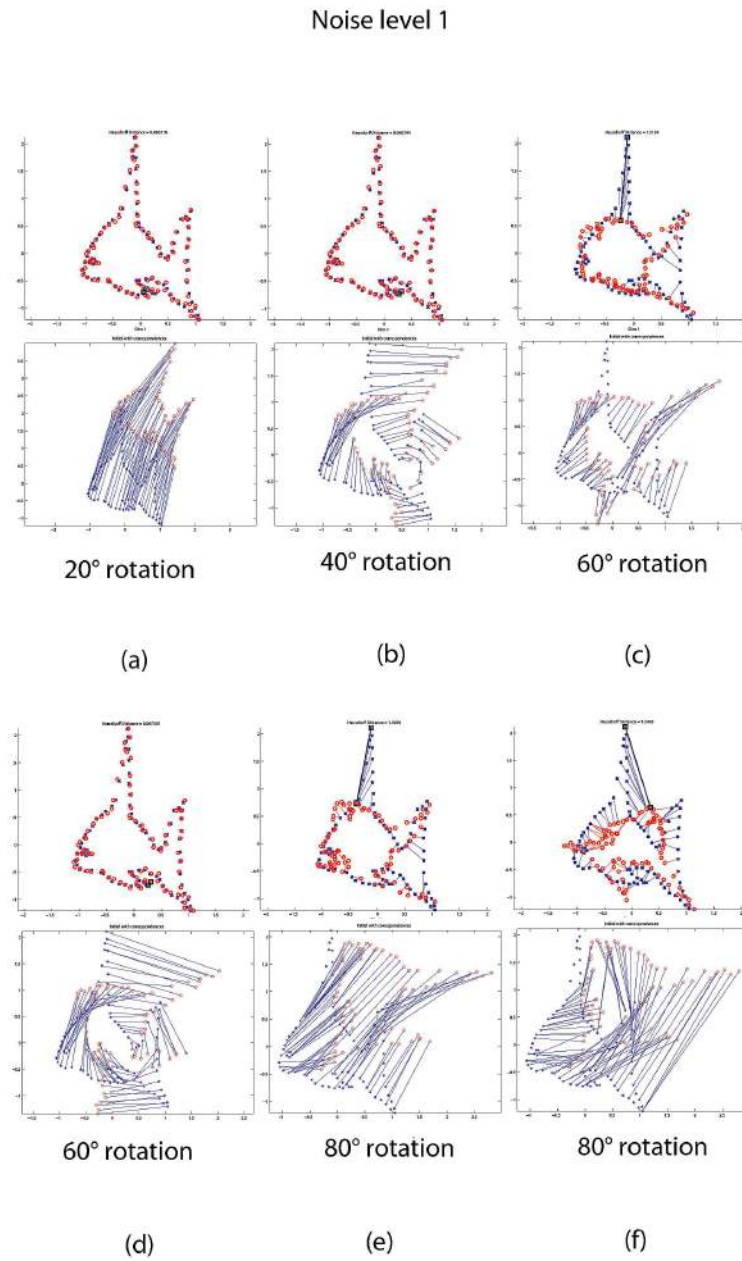


Figure 4.7: The algorithm [82] applied to the *sh* shape data set at noise level 1. (a), (b), (c), (e) represent the registration using the non rigid algorithm, (d), (f) the results using the rigid algorithm. Top row: the Hausdorff distance between the two shapes after the registration (black line between the black squares), also reported in the table 4.2. Bottom row: the fixed dataset is given by the blue stars, meanwhile the red circles represent the moving dataset. The lines give the correspondences between the fixed and the moving points. Observe that the correspondences in (c), (e), (f) are wrong (bottom row) and so is the registration in these cases (top row).

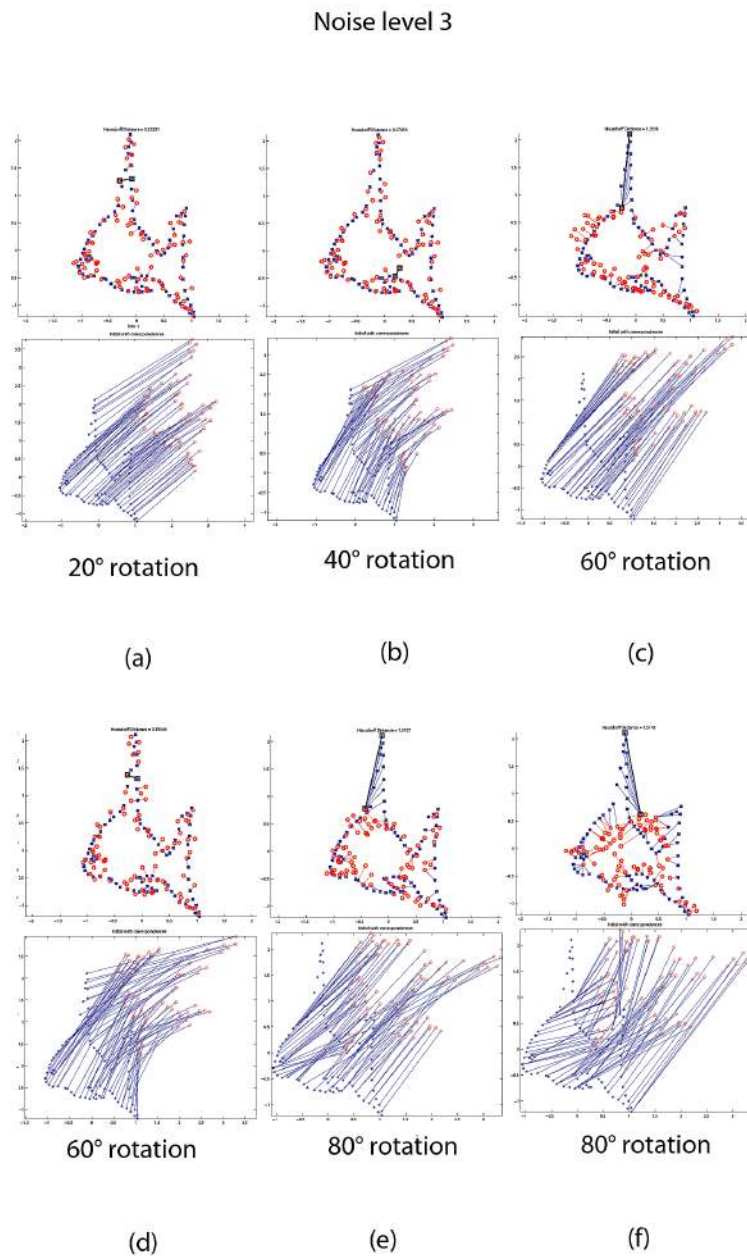


Figure 4.8: The algorithm [82] applied to the `sh` shape dataset at noise level 3. (a), (b), (c), (e) represent the registration using the non rigid algorithm, (d), (f) the results using the rigid algorithm. Top row: the Hausdorff distance between the two shapes after the registration (black line between the black squares), also reported in the table 4.2. Bottom row: the fixed dataset is given by the blue stars, meanwhile the red circles represent the moving dataset. The lines give the correspondences between the fixed and the moving points. Observe that the correspondences in (c), (e), (f) are wrong (bottom row) and so is the registration in these cases (top row).

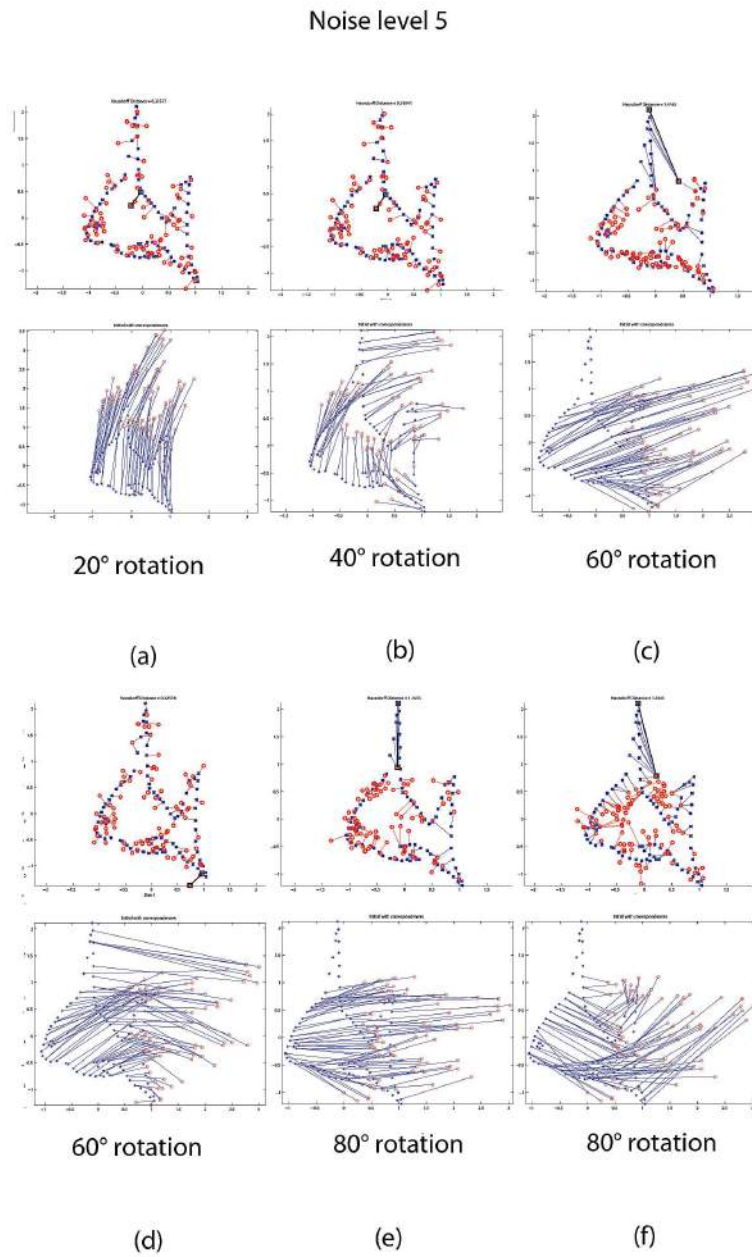


Figure 4.9: The algorithm [82] applied to the *sh* shape dataset at noise level 5. (a), (b), (c), (e) represent the registration using the non rigid algorithm, (d), (f) the results using the rigid algorithm. Top row: the Hausdorff distance between the two shapes after the registration (black line between the black squares), also reported in the table 4.2. Bottom row: the fixed dataset is given by the blue stars, meanwhile the red circles represent the moving dataset. The lines give the correspondences between the fixed and the moving points. Observe that the correspondences in (c), (e), (f) are wrong (bottom row) and so is the registration in these cases (top row).

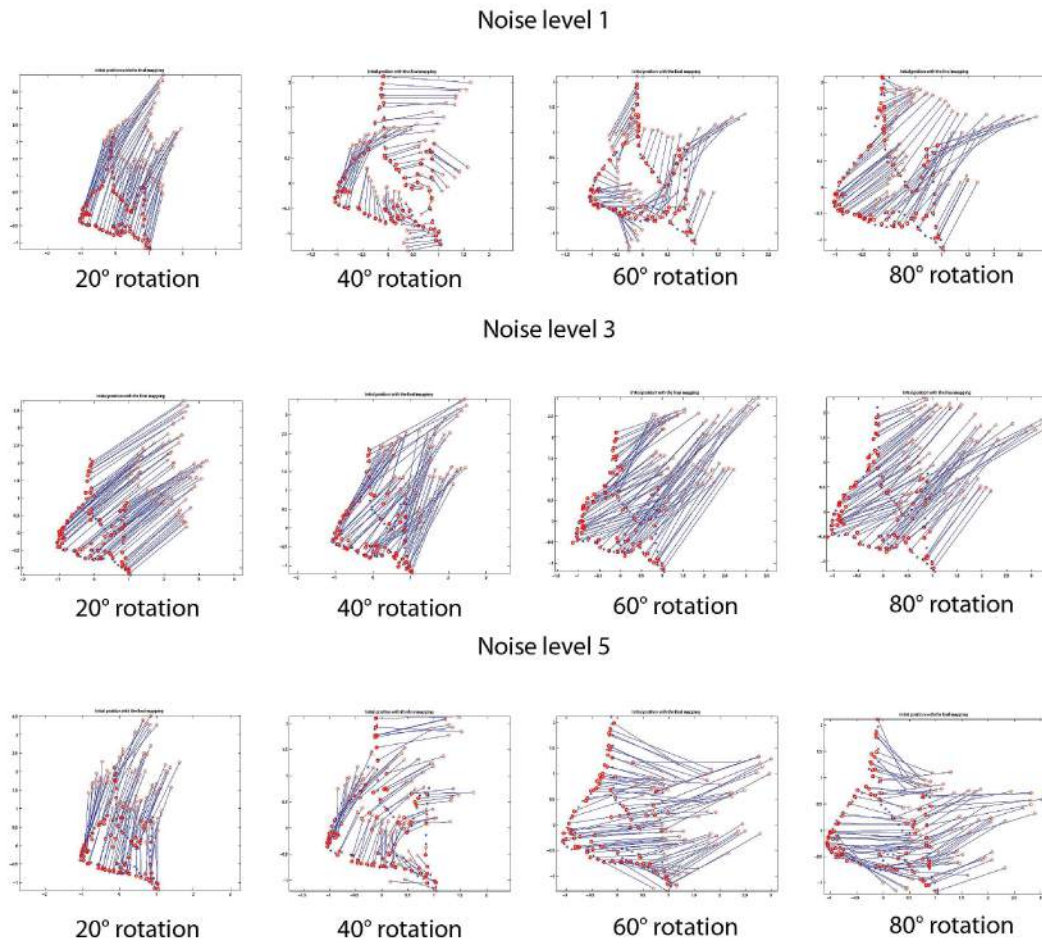


Figure 4.10: Tests applying the algorithm [53]. The fixed point are represented by blue stars, the moving points by thin red circles before the registration and by bold red circles after the non rigid registration. The lines give the movement of the correspondent red circles. The numerical values of the Hausdorff distance are given in the table 4.2.

4.6.5 Correspondences and registration evaluation using pulmonary landmark points derived from 4D CT image data

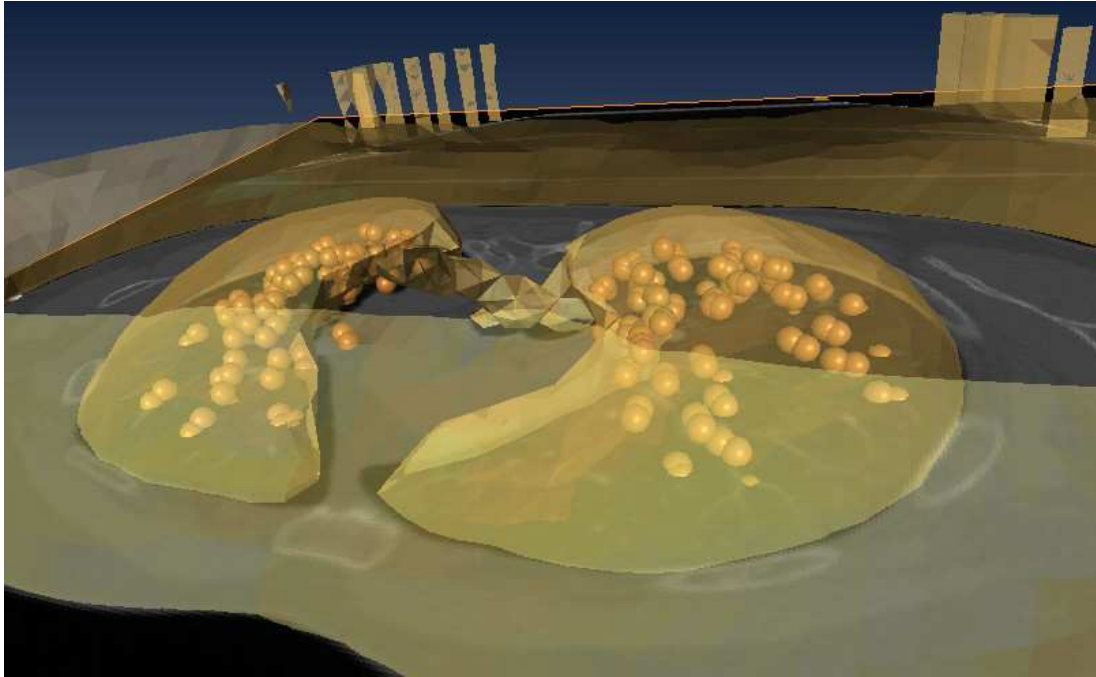


Figure 4.11: A 3D rendering of thoracic image with the pulmonary area and the extracted landmarks overlying a grayscale CT slice.

Thoracic 4D CT image data abound of high-contrast, anatomical landmarks such as vessel and bronchial bifurcations. Castillo et al. in [19] extracted a large number of landmark point pairs for the evaluation of deformable image registration spatial accuracy (Figure 4.11 and 4.12). We use this dataset to test the performance of our correspondence algorithm, then we register the landmarks using the rigid and non-rigid transformations.

The dataset from [19] includes 4D CT images from five patients free of pulmonary disease who were treated for esophageal cancer. Each patient underwent treatment planning in which 4D CT images of the entire thorax and upper abdomen were acquired at 2.5 mm spacing. In this study a high number of landmarks were manually extracted by an expert in thoracic imaging in 5 different breathing phases for each of the patient. The localization error of these landmarks is around 1mm.

For the purpose of our study we compared for each patient the initial breathing phase against the other 4 breathing phases by registering the landmarks without prior

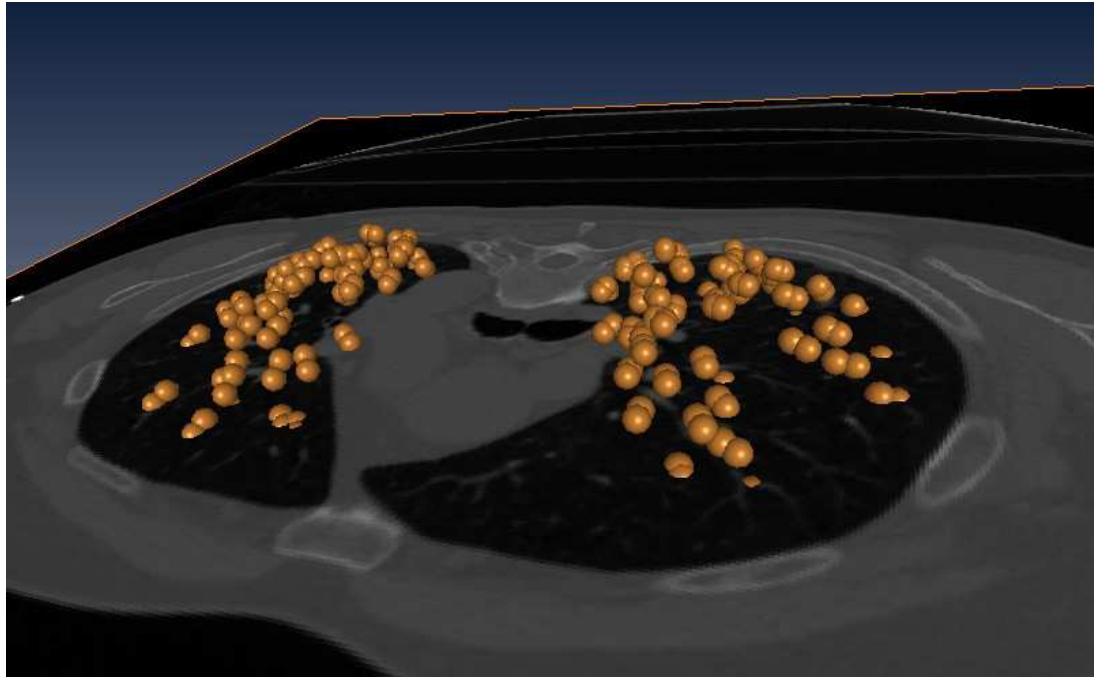


Figure 4.12: The 2D grayscale CT slice with 3D landmarks.

knowledge of the correspondences. After the computation of the landmarks correspondences, we registered the landmarks and we computed the Hausdorff distance, as in the previous section.

Table 4.3 summarizes the results of the correspondences expressed as a percentages.

Case number	Breathing phase			
	2	3	4	5
1	86%	80%	75%	60%
2	82%	79%	73%	59%
3	81%	76%	71%	58%
4	85%	80%	74%	59%
5	84%	78%	73%	57%

Table 4.3: Percentage of good correspondences for each patient and different breathing phases.

Estimates of the Hausdorff distance between the registered set of landmarks are summarized in table 4.4. The rigid registration is based on the correspondences found before. Since the number of good correspondences is greater than 50% of the total number of points, we have used an extraction of $p = 0.3$ of the best correspondences. For the non-rigid registration, we used the TPS interpolation as in the chapter 3, section

3.2. The input for the TPS interpolation is given by the same percentage of correspondences.

Case number	Breathing phase	Hausdorff distance (mm)	
		rigid	non-rigid
1	2	2.862	1.874
	3	3.583	2.936
	4	4.812	4.021
	5	6.104	5.513
2	2	2.729	1.908
	3	3.485	2.901
	4	4.764	4.177
	5	6.091	5.498
3	2	2.922	1.957
	3	3.579	2.990
	4	4.970	4.323
	5	6.245	5.556
4	2	2.820	2.015
	3	3.828	3.104
	4	5.057	4.238
	5	6.302	5.647
5	2	2.730	1.826
	3	3.749	2.857
	4	4.807	4.113
	5	5.970	5.324

Table 4.4: The Hausdorff distance after the rigid and the non rigid registration.

4.6.6 Conclusions and future work

The solution of the absolute orientation (Procrustes problem) is given by efficient methods in closed form, but to solve the registration of arbitrary point sets that represent the same shape acquired with different sensors and/or in different moments, there is no closed form solution. As summarized in the introduction, the solutions known so far range from iterative methods, where the optimization take place on each step toward a local minimum (as the ICP algorithm), capable to handle a certain amount of noise, to very complex methods that can handle also outliers and deformations, where the solution is given after a complicate process of optimization and using a large set of parameters.

In this chapter, we built the basis for a closed form solution, robust to a small

amount of noise, of the Procrustes problem in the case when no matching correspondences are given *a priori*.

The solution we have proposed makes use of the distribution of the distances. We have first analyzed, from a theoretical point of view, when the distribution of distances completely characterizes the shape of a point set and how to recover the isometric transformation between two sets of points, when the distribution of distances is given but the correspondences are not.

Beside this theoretical contribution, we have developed algorithms to find the correct alignment of shapes given as unstructured point sets. The registration took place after the correspondences were found and the iteration is used only to refine the distance between the aligned shapes.

We have seen in the beginning of this chapter that there are no efficient methods that can handle the variability of the data (i.e. noise, outliers, deformation) that can guarantee the goodness of the result and, in the same time, the possibility of the validation.

The methods to which we have compared our algorithm were registration algorithms that can theoretically handle deformations and outliers. We have shown that such algorithms fail even in recovering the global alignment transformation. The deformation model, used by these algorithms, allows to register by morphing, regardless of the fact that the shape may differ by noise or heavy rotations. These algorithms should be initialized with a good global registration algorithm, such as the algorithm presented here, to work properly.

Some of the research areas, where the registration requires a completely validated method, range from the medical applications (registration in the OR of surfaces extracted from the 3D reconstruction of organs) to robotic applications, where the accuracy and the real-time response are fundamental.

In the case of the medical applications the registration rely on datasets representing most of the time deformed surfaces but in order to map the instruments in the imaging space (e.g. biopsy needle, endoscopic camera, laparoscope etc.) we still need to isolate the rigid component of the registration.

With the solution we propose in this chapter, we are able to register noisy data represented by sets of points.

The test on medical data we have performed in this chapter employed a thoracic dataset from which a set of landmarks was extracted on different breathing phases.

The algorithm we have introduced was able to recover a high number of correspondences and to extract the rigid transformation between the point sets without having the homologies between the landmarks before. A further refinement was obtained by applying a non-rigid transformation to the result.

The next step will be to extend the method to handle incomplete data and outliers.

To replace the ICP algorithm with a closed form variant, in the case of dense point sets we may subsample to provide the candidates for potential correspondences. The main question is to subsample the points in a way that such that subsample points are still matchable. The joint clustering-matching algorithm presented by Chui in [23] can be a valid alternative for this extension.

If we replace the Euclidean distances in the shape matrix with geodesics we may handle also a larger class of isometries that include also the so called inelastic deformations, i.e. deformations that do not stretch or tear the object (see also [16]).

The results we have presented may be used also to extend the theory of shapes as defined by [58]. The equivalence class of k -ad points in \mathbb{R}^d will handle in this case also the permutations. It will be interesting to study how the topological properties of the new formed equivalence classes will change.

Chapter 5

A needle guidance system for percutaneous procedures

The progressive development of man is vitally dependent on invention.

Nikola Tesla

As stated in the beginning, the registration is not the end goal of our work, but we are interested to develop technologies to assist minimally invasive surgery. An example of a simple intervention is given in this chapter, together with the solution we have proposed. The imaging system is integrated in the procedure by the means of the registration. In this way the outcome of the procedure improves in terms of accuracy and time of execution, increasing the recovery time and the wellness of the patient.

In this chapter we introduce the design of a new navigation system for interventional radiology, implemented in a light and compact device. The system attached to the needle is composed by a small screen that gives indications about the position and the orientation of the needle, a controller that commands the screen and interfaces to the computer, and a marker that is identified by a tracking system (figure 5.1). By using a real time software the user is guided to move the needle along the desired position and orientation. To the best of our knowledges, this is the first system to have the navigation display integrated directly on the tool. Our in-vitro tests, show how such a

Based upon: D.Dall'Alba, B. Maris, P. Fiorini. *A compact navigation system for free hand needle placement in percutaneous procedures*, 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura, Portugal.

system yields a higher precision in the execution of the task than a free hand needle insertion and a reduction of the time required to complete the procedure.

5.1 Introduction

Clinical practice is increasingly replacing traditional open surgical procedures with minimally invasive techniques. This development results in a transition from direct visual feedback to image-based feedback. Many diagnostic and therapeutic procedures require high accuracy in placing an elongated instrument inside the target structure (e.g., biopsy, or ablative approaches) and they require great skills in operating the tools. Image-guided percutaneous procedures are used for both diagnosis and treatment. Diagnostic procedures include percutaneous biopsy (e.g. liver, kidney, pancreas, prostate biopsy) of tissues usually presumed to be tumoral. Therapeutic procedures include tumor ablation by different techniques such as radiofrequency ablation (RFA), where the dysfunctional tissues are ablated using high frequency alternate current, cryoablation (CA) where the treatment of the tumor is done by the insertion of a cryoprobe that freezes the surrounding tissues from the tip, or tumor embolization by the introduction of specific substances into the circulation to occlude vessels.

These procedures typically require precise insertion of an elongate instrument into the target organ. Image guided (IG) percutaneous ablation refers to the use of an imaging device (computed tomography or CT, magnetic resonance or MR, ultrasound or US) during the intervention to precisely place the ablation probes directly through the skin in order to destroy the tumor. The success of the ablation is constrained by the precise positioning of the needle. One of the techniques used to overcome this problem is to put a "targeting template" (paper with markers that can be seen with the CT or MRI) on the patient's back before imaging [7]. The needle based procedures could be subdivided in 3 main steps: the localization of the entry point on the skin of the patient, the orientation toward the target point and the needle insertion.

During a percutaneous procedure the mapping of the different information derived from the images acquired in the operating room is done mentally by the radiologist. While the radiologist inserts the needle, repeated series of CT (or MRI) scans are acquired to verify the correct insertion of the needle. When CT or MRI are used, the operation of needle insertion is performed blind, without any real time feedback.

As reported in [49], one of the most underestimated challenges of surgical guidance is the intraoperative display of surgical position and orientation during the case. The

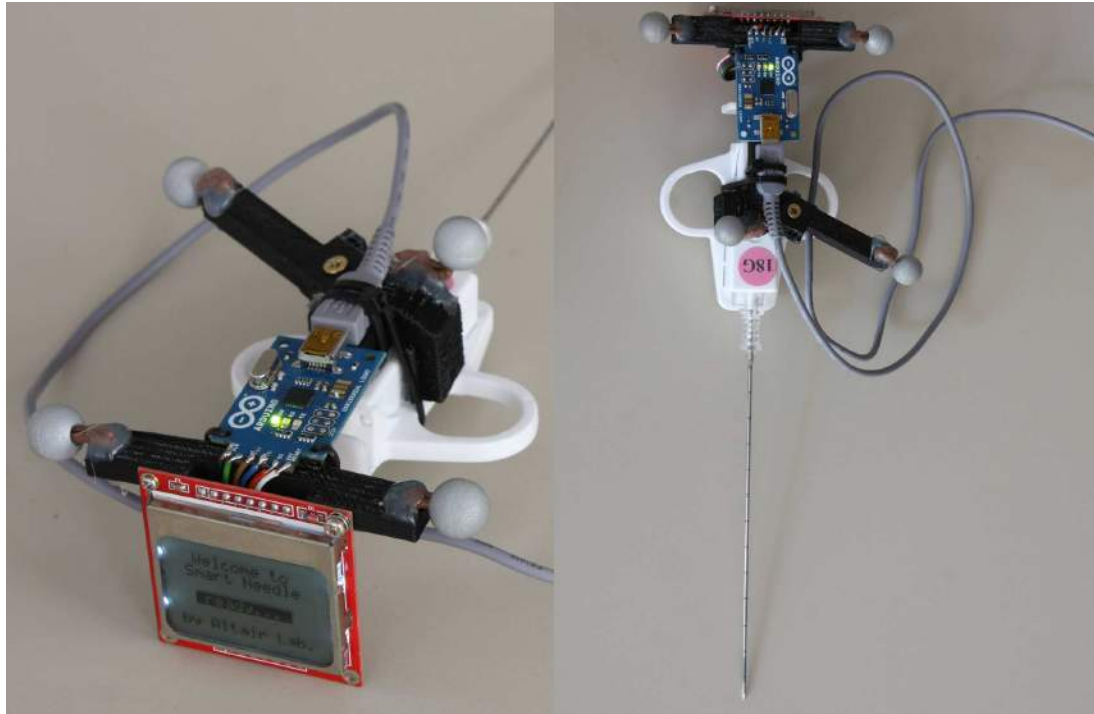


Figure 5.1: A standard 18 gauge biopsy needle equipped with the compact navigation system.

display is attempting to provide seven-dimensional information (X, Y, Z, Yaw, Pitch, Roll, and Time) on a two-dimensional, temporally active display. As the exact information needed cannot be displayed, the image-guided surgical displays are assistive and complement surgical knowledge and intuition.

5.2 Background

Numerous types of surgical navigation systems have been developed to aid the operator when tracking the surgical tool. The tracking is done by dynamically referencing ducials (skin markers, bone screws, a head frame, etc.) attached to the patient body.

The work done in [5] proposes a real-time electromagnetic position sensing of the needle tip to help the precision guidance into a liver for the radiofrequency ablation of metastatic tumors. A cross-hair targeting window on the computer screen allows the user to keep the needle's orientation correct during the puncture. A 2D image overlay device to assist needle placement on CT scanners is presented in [32]. The system demonstrated strong potential for reducing faulty needle insertion attempts,

thereby reducing X-ray dose and patient discomfort. In [59] the authors presented another augmented overlay imaging system integrated with an electromagnetic tracking system, which allows tracking of instruments, and of the radiologist's head position. The information is provided to the users through a transparent display. Electromagnetic tracking and the integration of CT and US images for needle placing in ablation procedures is described in [62]. A navigation system developed for minimally percutaneous interventions in the field of interventional radiology that can be connected to a CT scanner but also to C-arm systems is explored [84]. In [85] two jointly calibrated cameras are viewing the skin of the patient who is lying on the CT-table. A 3D model of the patient (including skin, liver and tumors) is automatically obtained from the CT scan, and is rigidly registered in the camera frame thanks to radio-opaque markers previously attached to the patients skin. The needle tracked by the cameras is displayed on a screen together with its position relative to the patient model. In [122] a magnetic-field-based navigation system is presented, which is an effective adjunct tool for accurate and safe biopsy of lesions that require an out-of-plane CT approach. In [120] the authors evaluate an augmented reality (AR) system in combination with a 1.5-T closed-bore magnetic resonance (MR) scanner as a navigation tool for needle biopsies. Tekbas [114] estimates the position of an abdominal target from a set of optically tracked fiducial needles. Prior to the intervention, the needles are inserted in the vicinity of the target, and a planning CT scan is acquired.

To overcome the difficulties of precise insertion of the needle, once the data acquired are in the same reference system, different approaches were proposed, for instance the use of a robotic arm [71].

All these works prove that the introduction of a navigation system during the procedure improve the performance in the needle insertion task. Anyway, no work addresses the specific problem of providing the navigation information to the user in an intuitive and ergonomic way.

The positioning of the needle is still a key issue and most of the time the percutaneous procedure is human based, therefore we propose here an integrated navigation system that provides information directly on the tool. The system helps the user to execute the procedure with high accuracy and precision.

5.3 System Work ow

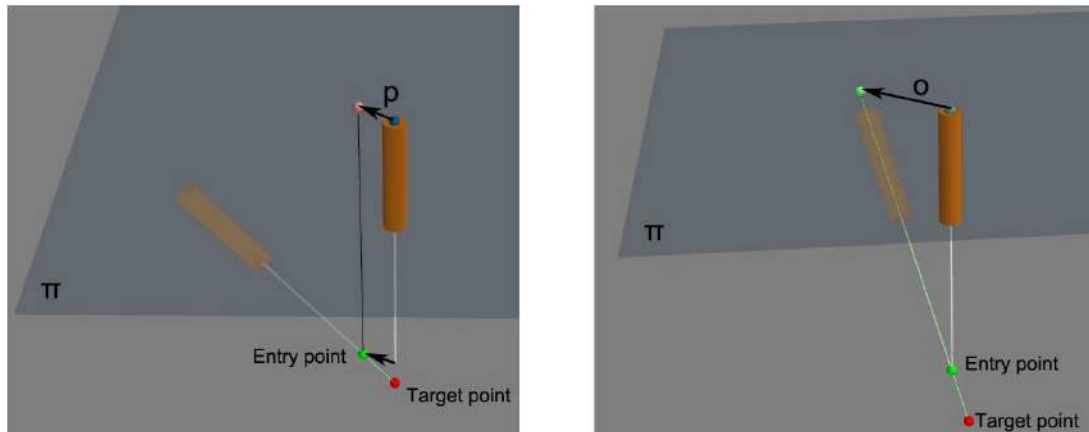


Figure 5.2: The needle navigation scheme. The plane π is perpendicular on the needle axis. The vector $\mathbf{p} \in \pi$ is the direction toward the entry point. The vector $\mathbf{o} \in \pi$ is the direction toward the correct orientation.

The proposed navigation system (Figure 5.1) is designed to be attached to any type of needle (the first prototype is designed for biopsy needles). This characteristic requires the precise identification of the needle geometry (calibration phase) together with a method to map the navigation information on a display attached to the tool (navigation phase).

The purpose of the calibration is to identify the needle reference system, where the needle axis represents the Oy -axis of this reference system, and to identify a direction in a plane π perpendicular to this axis (see Figure 5.2). This direction will be used as a reference, when mapping vectors on this plane. Once the calibration is performed, during the navigation we project the planned position and orientation onto the plane π .

The whole process is modeled as a state machine, where the initial state is the calibration, and the navigation is divided into entry point localization, finding the orientation and insertion (see Figure 5.3).

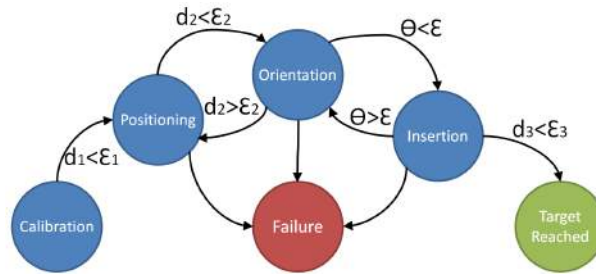


Figure 5.3: State machine diagram. The threshold error depends on the state: $\varepsilon_1 < 1mm$, $\varepsilon_2 < 3mm$, $\varepsilon_3 < 4mm$. The value of the ε_0 depends on the accuracy of the orientation we want to reach. The value of ε_3 depends on the accuracy we want to reach and should consider the calibration and the tracking errors.

5.3.1 Hardware Setup

The position and the orientation of the needle is tracked with an optical tracking system from NaturalPoint (Corvallis, OR, US) composed by 12 USB infrared (IR) cameras. The system uses IR reflecting markers and the software provided by the manufacturer to measure position and orientation of objects. The precision of the tracking system with object moving at speeds comparable with those of hand-held surgical tool is less than one millimeter. The 12 cameras are arranged in a truss like structure located over the operating table. This configuration ensures the line of sight between at least 2 cameras and each of the markers applied on the tool and does not interfere with the normal operations in the operating room.

To track the position of the needle is necessary to attach a rigid body to it. The rigid body is designed to guarantee that the display is orthogonal to the needle direction, therefore the navigation information are easily to be followed and to be visualized (figure 5.1). On this rigid body 4 infrared reflecting spheres were mounted in a unique geometric pattern to allow pose measurement from the tracking system. On the same structure we have also mounted the hardware components of the system, such as the display and the interface with the navigation workstation. A 2D Cartesian reference system of the display was defined with origin in the center of the display, and the axis parallel to the display sides. The x-axis is aligned also with the markers placed on the rigid body adjacent to it (see figure 5.4). This choice was selected to simplify the visualization of the navigation information through the display.

Once the needle is placed in the planned position, the marker structure can be removed to ease the sample withdrawal or the connection with the ablation device.

The display used in this first prototype is a 1.4" graphical mono-chrome display with a resolution of 84x48 pixel and a led back-lighting. This choice was done to reduce the power consumption of the hardware components in the design and also to increase the visibility of the display. In fact all the system is powered from a single USB port, the same used for communication.

5.3.2 Needle Calibration

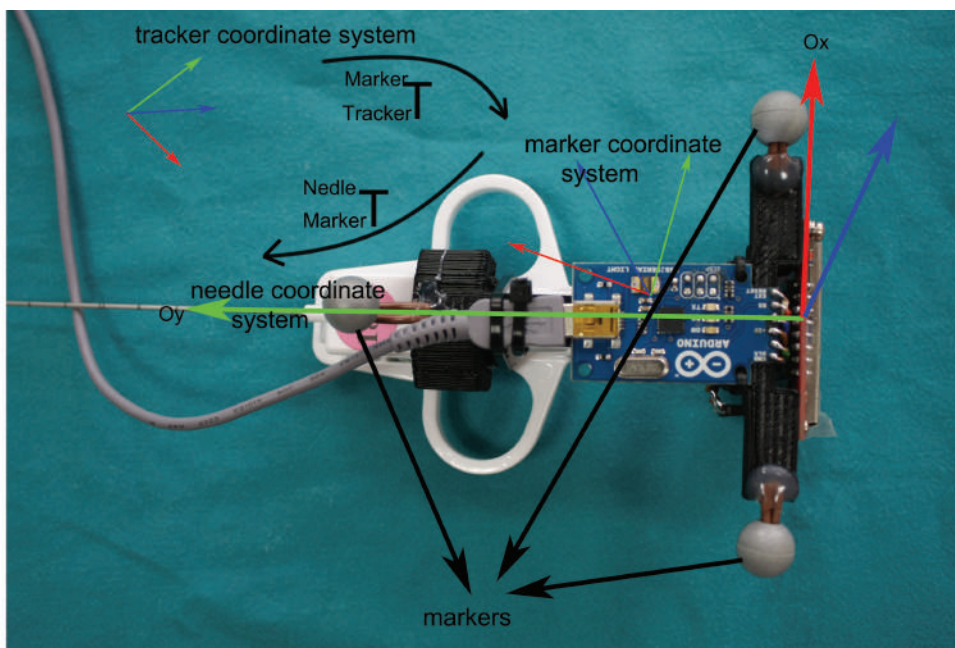


Figure 5.4: The coordinate systems rigid transformations. The matrix $\begin{matrix} Marker \\ Needle \end{matrix} T$ is the calibration matrix. The origin of the needle coordinate system is translated in the needle tip.

The pose of the needle is computed from the positions of the markers attached to it. In the first step of the calibration procedure, we identify the position of the needle tip relative to the markers by using the pivoting technique [125]. The orientation will be represented by the Oy -axis of the needle coordinate system and the tip position will give the origin of the reference frame (see Figure 5.4).

To compute the orientation of the needle, some systems use the insertion of markers along the needle axis [114]. The main drawback of these systems is the dependence on the accuracy of the tool design. The accuracy claimed in [70] for such a system is between 1 and 2 mm summed with a construction error of about 0.8 mm. Another

drawback of a tool that uses only marker along its axis, usually 2, is the loss of a degree of freedom (DoF).

To calibrate a needle without the application of markers to it, we should precisely identify two points along the needle axis. With this information we can align one axis of the needle reference frame along the axis of the needle.

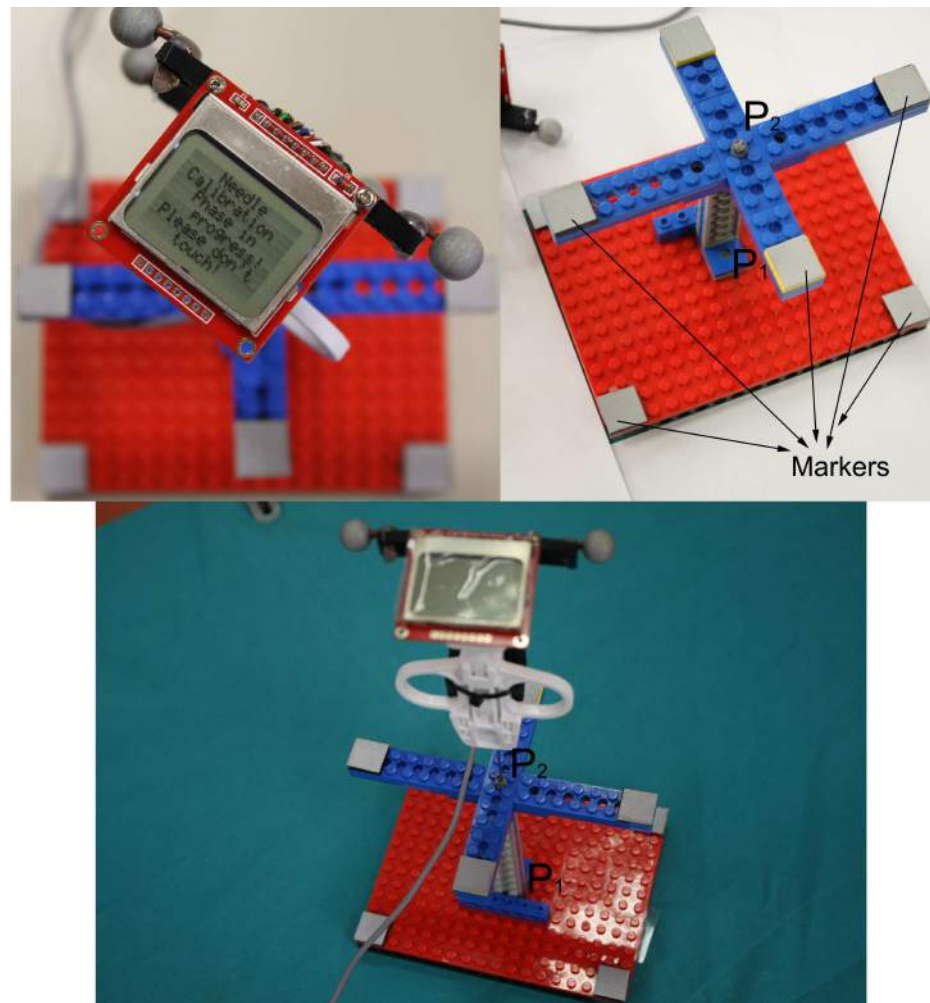


Figure 5.5: Cross calibration system.

We have designed and tested a cross-calibration system that helps to fully identify the position and the orientation of the needle. Our requirements were to keep the needle clean of markers and to have 6 DoF tracking. The calibration system consists of two rectangular plates, with 4 markers each, that precisely localize a point defined as the diagonals intersection. The system allows the needle to pass through these points (Figure 5.5). The automatic acquisition of the diagonal intersection was tested also by

pivoting and the results were essentially the same, the error is close to the error of the tracking system. In the operating room such a system allows a fast, easy and automatic procedure for the needle calibration. A first prototype was built with LEGO pieces, and the accuracy was lowered by this setup, but a steel model will be provided to ensure better robustness and the possibility of sterilization.

Once the needle is placed in the cross calibration system, a series of poses are acquired. To interpolate the rotation matrix we used quaternions and the SLERP method [29]. The interpolation of the translation vector is done by averaging the values of each translation vector. The two diagonal intersection P_1, P_2 represent one of the needle coordinate system axis (figure 5.5), where P_1 is the origin and P_2 is a point along one of the axis (e.g the y axis). The other axis of the needle coordinate system (Ox in the figure 5.4) is given by the projection of two markers on a plane perpendicular to the Oy -axis computed in the first step, therefore, after the calibration process, we have 6 DoF.

5.3.3 Needle Navigation

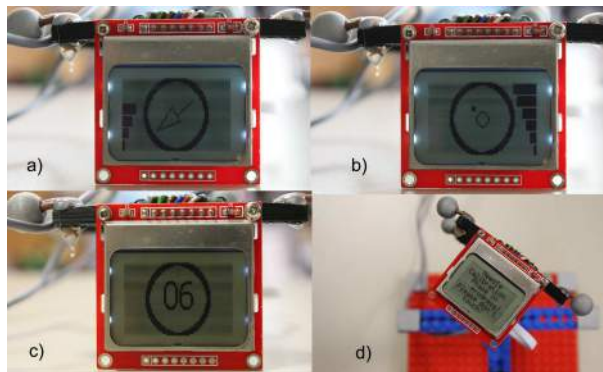


Figure 5.6: State machine display a) positioning state. b)orientation state. c) insertion step. d) calibration step.

For navigation purposes, once the needle is calibrated, we define a vector \mathbf{p} as the projection of the entry point to a plane perpendicular to the axis of the needle and the vector \mathbf{o} as the intersection of the direction between the line described by the entry point and the target point and the same plane (see Figure 5.2).

The first step of the navigation is the positioning of the needle at the entry point. The direction to be followed is the direction of the vector \mathbf{p} (figure 5.2). It is advantageous to perform the aligning step after finding the entry point, because the instrument

can be pivoted around the contact point between its tip portion and the surface without losing the entry point, which has already been located. During the second step of the navigation the needle is guided by the direction \mathbf{o} (figure 5.2). When the accurate orientation is found the insertion will take place. To pass from one state to another a threshold is defined depending on the accuracy we want to reach (see figure 5.3).

The control of the transitions between different steps of the procedure is implemented via software and the display is programmed to show the navigation information of each step of the state machine.

The 3D visualization shown on the PC monitor gives in real-time the position of the entry point, target point, the planned direction of the needle insertion and the actual needle position.

During the first step, i.e. the entry point position localization, the display attached to the needle provides the direction of the vector \mathbf{p} with an arrow and a bar chart that shows the distance to the entry point. The user has to move the tip of the needle parallel to the surface until the correct position is reached and the display shows the next phase (see figure 5.6 a).

In the orientation phase the user should keep the tip of the needle fixed until the desired direction is achieved. When the alignment is reached, the arrow indication is substituted with a small circle around the origin of the display that represents the target orientation, and a small point has to be moved inside this circle (see figure 5.6 b). The orientation angle error is computed as the cross product of the difference vector of the entry point and the target point and the vector along the needle axis. Once this value arrives below a threshold previously defined, the display shows the next step.

When the accurate orientation is found, the insertion step begins. In the first part of the insertion step the display shows the orientation information together with a bar chart that maps the distance from the target point. In fact, in a real percutaneous needle insertion, it is possible to change the needle orientation only during the first 20 mm of the insertion, therefore the necessity of having information on the orientation. In the last insertion part the display shows a distance indication in percentage (see figure 5.6 c).

Needle ID	1	2	3	All
Needle diameter	16G	18G	18G	
Position Analysis				
Accuracy RMS error (mean)	0.6561	0.7117	0.8002	0.7226
Accuracy RMS error (STD)	0.1674	0.3125	0.215	0.2316
Precision STD of recorded postions	0.3254	0.8613	0.9361	0.7076
Angular Analysis				
Accuracy [degree] (mean of dot angle)	0.3876	0.5372	0.8702	0.5983
Accuracy RMS error (STD)	0.2518	0.1872	0.3542	0.2644

Table 5.1: Results for the needle position and orientation calibration.

5.4 Validation and Results

5.4.1 Calibration Accuracy and Precision

An experimental study is conducted by using a grid matching approach as in [52], to assess the needle tracking accuracy and precision for both position and orientation. To evaluate the positioning accuracy a set of $n = 20$ points on the grid were identified first by pivoting in the tracker coordinate system P_T^1, \dots, P_T^n and then the same points were mapped to the tracking coordinate system using the calibration matrix and the positions in the needle coordinate system P_N^1, \dots, P_N^n .

$$P_i = {}^T_M T_{iN}^M T P_N^i \quad (5.1)$$

where ${}^T_M T_i$ is the tracking matrix of the markers on the needle and ${}^M_N T$ is the calibration matrix.

The orientation accuracy is given by:

$$\varepsilon_\theta = \text{dot}(P_k - P_l / \|P_k - P_l\|, O_y^T) \quad (5.2)$$

where P_k and P_l are two points crossed by the needle expressed in the tracker reference frame and

$$O_y^T = {}^T_M T_{iN}^M T [0100]^t \quad (5.3)$$

is the vector corresponding to the alignment of the needle Oy axis expressed in the tracker reference frame.

The root mean square (RMS) needle tip tracking error was defined as the RMS distance between the reference positions computed with the linear least square closed method and the transformed measured points:

$$\epsilon_{RMS}^{accuracy} = \sqrt{\frac{1}{n} \sum_{i=1}^n \|P_T^i - P_i\|_2^2} \quad (5.4)$$

The tracking precision for the needle tip P_j at the position j was defined as the RMS distance between the measured tip position and the mean tip position $P_{mean} = \frac{1}{n} \sum_{i=1}^n P_i$:

$$\epsilon_{RMS}^{precision} = \sqrt{\frac{1}{n} \sum_{i=1}^n \|P_i - P_{mean}\|_2^2} \quad (5.5)$$

where P_i are computed by eq. (5.1).

Table 5.1 reports the accuracy and the precision in positioning and orientation. The estimation was computed over 20 trials with 3 needles of different thickness expressed in gauge. The larger error was obtained with the thinner needles (18G) due to their flexion.

5.4.2 In-vitro evaluation of the navigation system

The performance of the developed device was tested in-vitro. A phantom similar to human skin was used, where the needle orientation is constrained by the skin and the muscular strip after a penetration of 1-2 cm. A simple polypropylene phantom with a thickness of 2 cm is well suited for this experiment since it allows the needle to keep the same orientation after insertion.

To localize a point inside the phantom we used a calibrated needle (target needle). The error in the localization of the target point is given therefore by the calibration error of the target needle. In the real case scenario the localization error is given by the error of the localization of the target. We x 5 entry positions on the surface of the phantom. The precise computation of the coordinates of the entry points is done as before, with the pivoting technique. Each of the entry point together with the target point defines an orientation.

For each of the 5 entry point positions we repeat a needle insertion 5 times, therefore we have a total of 25 insertions for each user.

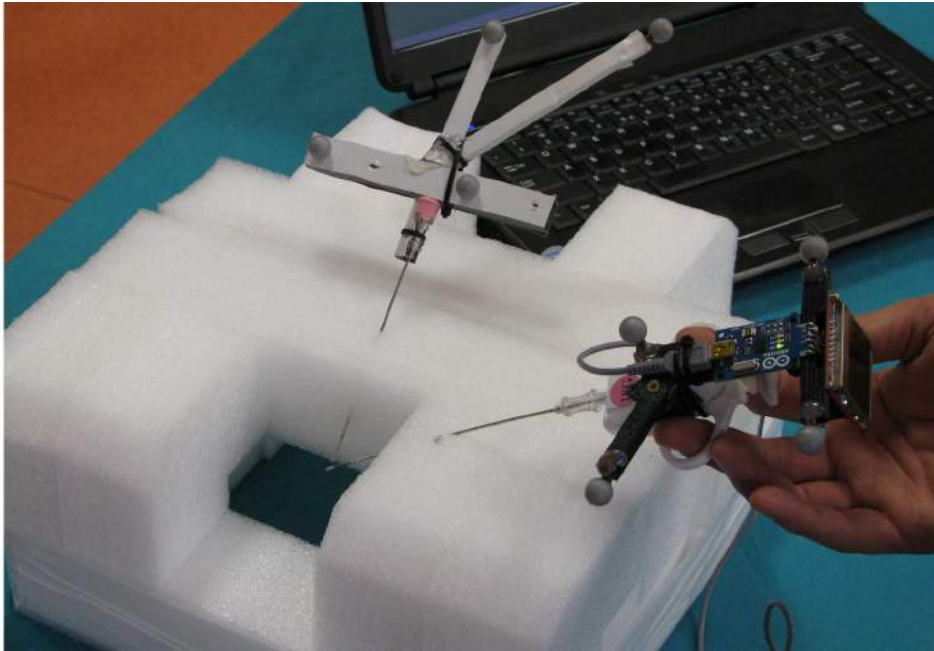


Figure 5.7: Experimental setup no. 2 with both the PC monitor and the display attached to the needle.

In the experimental setup no. 1 the needle insertion was performed with the use of a PC monitor placed in front of the user on the same table of the phantom. The 3D information about the tool and the points (entry and target) are represented on the monitor, as in a standard image guided procedure.

The experimental setup no. 2 was performed as the setup no. 1 but with the use of the new system that integrates the display applied on the needle (see Figure 5.7).

An important measure was the deviation from the desired orientation, once we have reached the target point. Since the planning phase defines the safe orientation, we have computed the angle between the planned orientation and the actual orientation. The positioning error was measured as the Euclidean distance between the needle tip position and the entry or the target point position, while the orientation error is measured as the dot product between the normalized vector represented by the entry point and target point and the vector representing the orientation of the needle given by the rotation matrix, since the needle axis is oriented with the coordinate system axis (see Figure 5.4).

Six subjects without previous experience tested the system after the oral explanation of the task. All the users are male and right hand dominant. We compared the results in terms of time taken to reach the goal and in terms of accuracy in entry and

target point localization.

Setup 1 with PC monitor						
	ELE [mm]		TLE [mm]		AOE [degree]	
User	Mean	STD	Mean	STD	Mean	STD
U1	5.114	1.453	4.50	2.542	5.839	1.667
U2	4.006	2.187	5.464	1.410	5.758	1.146
U3	5.081	2.321	4.506	2.187	6.259	2.869
U4	4.136	3.194	5.614	2.739	5.727	2.778
U5	5.147	1.492	4.497	1.665	5.420	1.726
U6	3.876	4.261	5.315	1.765	5.790	1.955
All	4.56	2.485	4.983	2.051	5.799	2.007
Setup 2 with PC monitor and display on the needle						
User	Mean	STD	Mean	STD	Mean	STD
U1	2.79	1.091	3.637	1.121	1.734	1.309
U2	2.703	1.114	3.079	1.677	2.456	1.856
U3	2.697	1.462	2.467	1.578	2.902	1.068
U4	2.727	1.257	3.916	1.988	1.833	1.576
U5	2.916	1.725	2.991	2.091	2.52	1.848
U6	2.627	2.111	3.043	1.26	4.036	0.918
All	2.746	1.46	3.196	1.619	2.58	1.429

Table 5.2: Comparison of Error Values in the two experimental setups.

Experimental Setup 1			Experimental Setup 2	
	Time [s]		Time [s]	
User	Mean	STD	Mean	STD
U1	127.11	22.35	90.2	10.65
U2	137.23	19.43	73.64	26.83
U3	128.31	35.23	77.11	18.23
U4	144.94	11.82	94.58	16.73
U5	135.97	14.4	82.48	15.33
U6	129.29	29.33	70.81	15.95
All	133.78	22.09	81.47	17.29

Table 5.3: Comparison of Timing in the two experimental setups.

In table 5.2 we report the accuracy and precision results of the experiment. We show the entry point localization error (ELE), the target point localization error (TLE) and angular orientation error (AOE). The results of experimental setup 2 show that the mean and STD of the error is lower than in setup 1 for all the errors considered.

When using the proposed approach, the entry position and the orientation of the needle must be below the threshold imposed by the system to pass from a state to another (see figure 5.3). The error in the target positioning will then be constrained by the errors accumulated in the system and by the calibration errors.

The use of the proposed system reduces the error (both in the position and in the orientation) and also reduces the influence of the user's experience in needle insertion execution. In fact the standard deviation of the data is lower and more stable than in the case without display on the needle.

The needle insertion time, measured from the beginning of the localization of the entry point position, with the display mounted on the needle is about 40% lower than without the display (see table 5.3). The strong improvement in the performance is probably due to the lack of experience of the subjects and in some limitations of the design of the experimental setup.

5.5 Conclusions and Future Work

This chapter proposed a new system for accurate needle placement. We have designed, tested and evaluated the system in an in-vitro experiment that was not directly related to a clinical scenario. The experiments we have done validate the utility of the tool that can be used in a more complex procedure.

The main advantages of the navigation system we have presented are:

- gives the information directly on the tool.
- there is no need for the user to move the eyes from the interventional area.
- is fast and simple to calibrate, therefore easy to be integrated in the operating room scenario.
- improves the accuracy and the precision to reach the desired position and orientation.
- reduces the time needed to complete the procedure, gives the possibility to map the needle in an environment where other tools are present by the mean of the tracking system.(e.g. US scanner, 3D surface scanner).

In the future works the electronic components will be miniaturized, and also the display should have a higher resolution and a higher refresh rate. The rigid body

design will provide more protection to the electronic components. Both the calibration phantom and the needle device are being redesigned to take into account cleaning and sterilization requirements imposed by the operation room environment.

The performance of the navigation system should be evaluated with users that have previous background in percutaneous procedures.

The navigation system will be included in a percutaneous image-guided procedure, where pre-operative images are available and intra-operative real-time images will be acquired and registered to the same reference system.

Chapter 6

Final words and future works

... mathematics may be defined as the subject in which we never know what we are talking about, nor whether what we are saying is true.

Bertrand Russell

In this thesis an analysis of the registration algorithms to be applied in the OR was performed. Since there is a lot of work in this area, we have identified the direction to proceed in order to find a solution for our applications. We have seen that the main theory in the field of medical image registration is not usable in the OR (chapter 2). On the other hand, the techniques already in use in the OR suffer from many limitations. They are based on the manual extraction of landmark followed by a rigid registration (chapter 3).

The only well accepted and employed algorithm, ICP, has the same limitations of the other algorithms that use optimization and distance measures: the impossibility to avoid local minima and the lack of control over the iteration limit.

We have performed an in-vitro test of the principal-axes registration in order to find alternative approaches to the existing ones (chapter 3). The solution we have used worked well but the assumptions we have made were strong. The requirements of a registration algorithm for the OR were therefore identified :

- complete integration in a larger system that comprises imaging devices both for acquisition and visualization, tracking tools, medical instruments.
- accuracy and precision.

- validation of the results, i.e. the possibility to have control over the registration results in order to estimate beforehand the precision of the overall system.
- real-time performance.

The algorithms we proposed in chapter 4 are oriented toward these goals.

We developed our approach starting from the closed form solution of the Procrustes method. The first solution we gave, algorithm 1, holds in the ideal case, when no noise or little noise is present, and yields the transformation that registers one set of points to another, without knowing the correspondences. The second solution, algorithms 2 and 3, solves the correspondences in the ideal case and in the noisy case respectively, and derives the transformation that register one set of point to another from the computed correspondences (section 4.6.4).

Finally, in the chapter 5, a system for needle navigation to perform minimally invasive manual technique was introduced. The registration in the OR is not a process on its own but is part of larger systems where the integration of different data is of paramount importance for the precision of the intervention and the safety of the patient. The minimally invasive techniques are sometimes guided blindly (e.g. in the prostate biopsy the tumor cannot be seen with US real-time imaging) and the design of new tools together with the integration of registration can lead to substantial improvements of the interventions.

Most of the results we have presented here were obtained on synthetic data acquired from anatomical phantoms. This choice was motivated on the one hand by the need to have a controlled environment, in order to measure the errors introduced by the system and, on the other hand, by the fact that most of the work in this thesis derives from the interaction of the individual contribution with international research projects that deals with minimally invasive procedures and surgical robotic procedures that are still ongoing research and must be validated before entering the OR. Nevertheless, our work was guided by the feed-back of many physicians and some of the techniques were developed after attending real-life interventions.

The next step will be the effective introduction of the technologies we developed so far in the clinical practice by a strong collaboration of both researchers and physicians.

Nowadays the technology is evolving and we have the conviction that there will be a large use of it in the very near future also in the clinical context, in spite of the fact that the validation and the bureaucratic processes are not easy.

As far as the registration of medical images in the OR concerns, even though there is a very high demand in the OR, we think we are still at the beginning and this thesis

gave a contribution not only to find the solution to this problem but also to point the research to the right direction.

We plan to extend the theory we have developed in chapter 4 to handle inelastic deformations in order to have not only the rigid correspondence but also the correspondence generated by this type of deformation. This step is not difficult to achieve, the only important issue to address is the efficient computation of the geodesics for the real-time acquired data. Once the correspondences are defined, a non-rigid model, e.g. given by spline functions, will yield the desired transformation.

Another line of research in which we are interested in the future is the integration of the registration in larger systems that assist OR interventions. This will be achieved by the introduction of new imaging systems in the OR and by the development of new tools.

Bibliography

- [1] N M Alpert, J F Bradshaw, D Kennedy, and J A Correia. The principal axes transformation—a method for image registration. *Journal of nuclear medicine of cial publication Society of Nuclear Medicine* , 31(10):1717–1722, 1990.
- [2] K S Arun, T S Huang, and S D Blostein. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.
- [3] Brian B. Avants, Charles L. Epstein, Murray Grossman, and James C. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, (1):26–41.
- [4] Kwang-Ho Bae and Derek D Lichti. Automated registration of unorganized point clouds from terrestrial laser scanners. *Update*, 35(5B):36–54, 2004.
- [5] Filip Banovac, Emmanuel Wilson, Hui Zhang, and Kevin Cleary. Needle biopsy of anatomically unfavorable liver lesions with an electromagnetic navigation assist device in a computed tomography environment. *Journal of vascular and interventional radiology JVIR*, 17(10):1671–1675, 2006.
- [6] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12:43–77, 1994.
- [7] J G Baust, A A Gage, D Klossner, D Clarke, R Miller, J Cohen, A Katz, T Polascik, H Clarke, and J M Baust. Issues critical to the successful application of cryosurgical ablation of the prostate. *Technology in cancer research treatment*, 6(2):97–109, 2007.
- [8] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *In ECCV*, pages 404–417, 2006.

- [9] S Belongie, J Malik, and J Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *In NIPS*, volume 54, pages 831–837. Cite-seer, Citeseer, 2000.
- [10] P J Besl and N D McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [11] F L Bookstein. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.
- [12] Mireille Boutin and Gregor Kemper. Which point configurations are determined by the distribution of their pairwise distances? *CoRR*, math.MG/0311004, 2003.
- [13] Mireille Boutin and Gregor Kemper. Which point configurations are determined by the distribution of their pairwise distances? *Int. J. Comput. Geometry Appl.*, 17(1):31–44, 2007.
- [14] Doug M Boyer, Yaron Lipman, Elizabeth St Clair, Jesus Puente, Biren A Patel, Thomas Funkhouser, Jukka Jernvall, and Ingrid Daubechies. Algorithms to automatically quantify the geometric similarity of anatomical surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, 108(45):18221–18226, 2011.
- [15] Chaim Broit. Optimal registration of deformed images. *Ph.D. dissertation, Univ. Pennsylvania, Philadelphia*, 1981.
- [16] Alexander Bronstein, Michael Bronstein, and Ron Kimmel. *Numerical Geometry of Non-Rigid Shapes*. Springer Publishing Company, Incorporated, 1 edition, 2008.
- [17] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.
- [18] M M Burlew, E L Madsen, J A Zagzebski, R A Banjavic, and S W Sum. A new ultrasound tissue-equivalent material. *Radiology*, 134(2):517–520, 1980.
- [19] Richard Castillo, Edward Castillo, Rudy Guerra, Valen E Johnson, Travis McPhail, Amit K Garg, and Thomas Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Physics in Medicine and Biology*, 54(7):1849, 2009.

- [20] Warren Cheung and Ghassan Hamarneh. n-sift: n-dimensional scale invariant feature transform for matching medical images. In *In Proceedings of the Fourth IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2007 (ISBI 2007)*, pages 720–723, 2007.
- [21] Gary E. Christensen and Hans J. Johnson. Consistent image registration. *IEEE Trans. Med. Imaging*, 20(7):568–582, 2001.
- [22] Gary E. Christensen, Richard D. Rabbitt, and Michael I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, 1996.
- [23] Haili Chui. *Non-Rigid Point Matching: Algorithms, Extensions and Applications*. PhD thesis, Yale University, May 2001.
- [24] Haili Chui and Anand Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.
- [25] Logan W. Clements, Prashanth Dumpuri, William C. Chapman, Benoit M. Dawant, Robert L. Galloway, and Michael I. Miga. Organ surface deformation measurement and analysis in open hepatic surgery: Method and preliminary results from 12 clinical cases. *IEEE Trans. Biomed. Engineering*, 58(8):2280–2289, 2011.
- [26] A Collignon, F Maes, D Delaere, D Vandermeulen, P Suetens, and G Marchal. Automated multimodality medical image registration using information theory. In Y Bizais and C Barillot, editors, *Information Processing in Medical Imaging*, pages 263–274. Kluwer Academic Publishers, Dordrecht, 1995.
- [27] Timothy F. Cootes, Carole J. Twining, Vladimir S. Petrovic, Kolawole O. Babalola, and Christopher J. Taylor. Computing accurate correspondences across groups of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(11):1994–2005, 2010.
- [28] Emiliano D'Agostino, Jan Modersitzki, Frederik Maes, Dirk Vandermeulen, Bernd Fischer, and Paul Suetens. Free-form registration using mutual information and curvature regularization. In *WBIR*, pages 11–20, 2003.

- [29] E B Dam, M Koch, and M Lillholm. Quaternions, interpolation and animation. *Spring*, page 103, 1998.
- [30] Jérôme Declerck, Jacques Feldmar, M. L. Gortis, and Fabienne Betting. Automatic registration and alignment on a template of cardiac stress and rest reoriented spect images. *IEEE Trans. Med. Imaging*, 16(6):727–737, 1997.
- [31] D W Eggert, A Lorusso, and R B Fisher. Estimating 3-D rigid body transformations: a comparison of four major algorithms. *Machine Vision and Applications*, 9(5-6):272–290, 1997.
- [32] Gabor Fichtinger, Anton Deguet, Ken Masamune, Emese Balogh, Gregory S Fischer, Herve Mathieu, Russell H Taylor, S James Zinreich, and Laura M Fayad. Image overlay guidance for needle insertion in CT scanner. *IEEE Transactions on Biomedical Engineering*, 52(8):1415–1424, 2005.
- [33] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [34] J. Michael Fitzpatrick and Jay B. West. The distribution of target registration error in rigid-body, point-based registration. *IEEE Trans. Med. Imaging*, 20(9):917–927, 2001.
- [35] Natasha Gelfand. Robust Global Registration. *Proceedings of the third Eurographics symposium on Geometry processing*, 38(2):197–es, 2005.
- [36] Deboshmita Ghosh, Andrei Sharf, and Nina Amenta. Feature-driven deformation for dense correspondence. pages 726136–726136–10, 2009.
- [37] A. A. Goshtasby. *Image Registration: Principles, Tools and Methods*. Springer, 2012.
- [38] Ardeshir Goshtasby. *2D and 3D image registration for medical, remote sensing and industrial applications*. Published by Wiley-IEEE, 2005.
- [39] Christoph Guetter, Chenyang Xu, Frank Sauer, and Joachim Hornegger. Learning based non-rigid multi-modal image registration using kullback-leibler divergence. In *MICCAI (2)*, pages 255–262, 2005.

- [40] Yujun Guo, Priya N Werahera, Ramkrishnan Narayanan, Lu Li, Dinesh Kumar, E David Crawford, and Jasjit S Suri. Image registration accuracy of a 3-dimensional transrectal ultrasound-guided prostate biopsy system. *Journal of ultrasound in medicine : official journal of the American Institute of Ultrasound in Medicine*, 28(11):1561–1568, 2009.
- [41] Eldad Haber and Jan Modersitzki. Intensity gradient based registration and fusion of multi-modal images. In *MICCAI (2)*, pages 726–733, 2006.
- [42] Jacques Hadamard. Sur les problemes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.
- [43] J V Hajnal, N Saeed, E J Soar, A Oatridge, I R Young, and G M Bydder. A registration and interpolation procedure for subvoxel matching of serially acquired MR images. *Journal Of Computer Assisted Tomography*, 19(2):289–296, 1995.
- [44] Joseph V. Hajnal, Derek L.G. Hill, and David J. Hawkes. *Medical Image Registration*. CRC Press, 2001.
- [45] T J Hall, M Bilgen, M F Insana, and T A Krouskop. Phantom materials for elastography, 1997.
- [46] Xiaohui Hao Xiaohui Hao, C Bruce, C Pislaru, and J F Greenleaf. A novel region growing method for segmenting ultrasound images, 2000.
- [47] Chris Harris and Mike Stephens. A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [48] Mattias P. Heinrich, Mark Jenkinson, Manav Bhushan, Tahreema N. Matin, Fergus Gleeson, Michael Brady, and Julia A. Schnabel. Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012.
- [49] S Duke Herrell, Robert L Galloway, and Li-Ming Su. Image-guided robotic surgery: update on research and potential applications in urologic surgery. *Current Opinion in Urology*, 22(1):47–54, 2012.
- [50] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

- [51] William M. Hsu, John F. Hughes, and Henry Kaufman. Direct manipulation of free-form deformations. In *SIGGRAPH*, pages 177–184, 1992.
- [52] Johann B Hummel, Michael R Bax, Michael L Figl, Yan Kang, Calvin Maurer, Wolfgang W Birkfellner, Helmar Bergmann, and Ramin Shahidi. Design and application of an assessment protocol for electromagnetic tracking systems. *Medical Physics*, 32(7):2371, 2005.
- [53] Bing Jian and Baba C Vemuri. Robust Point Set Registration Using Gaussian Mixture Models, 2011.
- [54] A.E. Johnson and Sing Bing Kang. Registration and integration of textured 3-d data. In *3-D Digital Imaging and Modeling, 1997. Proceedings., International Conference on Recent Advances in*, pages 234 –241, may 1997.
- [55] K Kanatani. *Geometric Computation for Machine Vision*, volume 37 of *The Oxford Engineering Science Series*. Oxford University Press, 1993.
- [56] Yan Ke and Rahul Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. In *CVPR (2)*, pages 506–513, 2004.
- [57] D G Kendall. Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984.
- [58] D G Kendall, D Barden, T K Carne, and H Le. *Shape and Shape Theory*. Wiley, 1999.
- [59] M Fawad Khan, Selami Dogan, Adel Maataoui, Stefan Wesarg, Jessen Gurrung, Hanns Ackermann, Mirko Schiemann, Gerhard Wimmer-Greinecker, and Thomas J Vogl. Navigation-based needle puncture of a cadaver using a hybrid tracking navigational system. *Investigative Radiology*, 41(10):713–720, 2006.
- [60] Jeongtae Kim and Jeffrey A. Fessler. Intensity-based image registration using robust correlation coefficients. *IEEE Trans. Med. Imaging*, (11):1430–1444.
- [61] Stefan Klein, Marius Staring, and Josien P. W. Pluim. Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines. *IEEE Transactions on Image Processing*, 16(12):2879–2890, 2007.

- [62] Jochen Krücker, Sheng Xu, Neil Glossop, Anand Viswanathan, Jörn Borgert, Heinrich Schulz, and Bradford J Wood. Electromagnetic tracking for thermal ablation and biopsy guidance: clinical evaluation of spatial accuracy., 2007.
- [63] Jan Kybic and Michael Unser. Fast parametric elastic image registration. *IEEE Transactions on Image Processing*, 12(11):1427–1442, 2003.
- [64] Paul Lemke, Steven S. Skiena, and Warren D. Smith. Reconstructing sets from interpoint distances, 2003.
- [65] Rui Liao, Christoph Guetter, Chenyang Xu, Yiyong Sun, Ali Khamene, and Frank Sauer. Learning-based 2d/3d rigid registration using jensen-shannon divergence for image-guided surgery. In *MIAR*, pages 228–235, 2006.
- [66] Yaron Lipman and Thomas Funkhouser. Mobius voting for surface correspondence. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 28(3), August 2009.
- [67] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.
- [68] E L Madsen, J A Zagzebski, R A Banjavie, and R E Jutila. Tissue mimicking materials for ultrasound phantoms. *Medical Physics*, 5(5):391–394, 1978.
- [69] F Maes, A Collignon, D Vandermeulen, G Marchal, and P Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
- [70] Lena Maier-Hein, Aysun Tekbas, Alexander Seitel, Frank Pianka, Sascha A Müller, Stefanie Satzl, Simone Schawo, Boris Radeleff, Ralf Tetzlaff, Alfred M Franz, Beat P Müller-Stich, I Wolf, H-U Kauczor, B M Schmied, and H-P Meinzer. In vivo accuracy assessment of a needle-based navigation system for CT-guided radiofrequency ablation of the liver. *Medical Physics*, 35(12):5385–5396, 2008.
- [71] Lena Maier-Hein, Conor J Walsh, Alexander Seitel, and Nevan C Hanumara. Human vs . robot operator error in a needle-based navigation system for percutaneous liver interventions Citation Accessed Citable Link Detailed Terms

- Human vs . robot operator error in a needle-based navigation system for percutaneous liver interventions. *Proceedings of SPIE*, pages 72610Y–72610Y–12, 2009.
- [72] J B Maintz and M A Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.
- [73] J. B. Antoine Maintz, Petra A. van den Elsen, and Max A. Viergever. 3d multimodality medical image registration using morphological tools. *Image Vision Comput.*, (1-2):53–62.
- [74] Siddharth Manay, Byung-Woo Hong, Anthony J Yezzi, and S Soatto. Integral invariant signatures. *Computer VisionECCV 2004*, pages 87–99, 2004.
- [75] Facundo Mémoli and Guillermo Sapiro. A Theoretical and Computational Framework for Isometry Invariant Recognition of Point Cloud Data. *Foundations of Computational Mathematics*, 5(3):313–347, 2005.
- [76] N J Mitra. Dynamic Geometry Registration. *Processing*, pages:182, 2007.
- [77] J. Modersitzki. *Numerical Methods for Image Registration*. Oxford University Press, 2004.
- [78] J. Modersitzki. *FAIR: Flexible Algorithms for Image Registration*. SIAM, Philadelphia, 2009.
- [79] Jan Moderstitzki and Bernd Fischer. Ill-posed medicine an introduction to image registration. *Inverse Problems*, 24(3).
- [80] Mehdi Hedjazi Moghari and Purang Abolmaesumi. Distribution of target registration error for anisotropic and inhomogeneous ducial lo calization error. *IEEE Trans. Med. Imaging*, 28(6):799–813, 2009.
- [81] Jean-Michel Morel and Guoshen Yu. Asift: A new framework for fully af ne invariant image comparison. *SIAM J. Imaging Sciences*, 2(2):438–469, 2009.
- [82] A Myronenko and Xubo Song Xubo Song. Point Set Registration: Coherent Point Drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–75, 2010.

- [83] Andriy Myronenko and Xubo B. Song. Intensity-based image registration by minimizing residual complexity. *IEEE Trans. Med. Imaging*, 29(11):1882–1891, 2010.
- [84] Markus Nagel, Martin Hoheisel, Ralf Petzold, Willi A Kalender, and Ulrich H W Krause. Needle and catheter navigation using electromagnetic tracking for computer-assisted C-arm CT interventions. *Proceedings of SPIE*, 6509:65090J–65090J–9, 2007.
- [85] S Nicolau, X Pennec, L Soler, and N Ayache. An augmented reality system to guide liver punctures. In *Workshop AMIARCS*, 2004.
- [86] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer, 1999.
- [87] Andreas Nuchter, Kai Lingemann, and Joachim Hertzberg. Cached k-d tree search for ICP algorithms. *Sixth International Conference on 3D Digital Imaging and Modeling 3DIM 2007*, 7(x):419–426, 2007.
- [88] Nobuyuki Otsu. A Threshold Selection Method from Gray-level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.
- [89] Gaurav Pandey, James R. McBride, Silvio Savarese, and Ryan M. Eustice. Visually bootstrapped generalized icp. In *Proceedings of the IEEE International Conference on Robotics and Automation*, Shanghai, China, 2011. Accepted, To Appear.
- [90] M Peterhans, A Vom Berg, B Dagon, D Inderbitzin, C Baur, D Candinas, and S Weber. A navigation system for open liver surgery: design, work ow and rst clinical applications. *Int J Med Robot*, 7(1):7–16, 2011.
- [91] Josien P. W. Pluim, J. B. Antoine Maintz, and Max A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Trans. Med. Imaging*, 19(8):809–814, 2000.
- [92] William H Press, S A Teukolsky, William T Vetterling, and Brian P Flannery. *Numerical Recipes*. Cambridge University Press, 2007.
- [93] Anand Rangarajan, Haili Chui, and Fred L Bookstein. The Softassign Procrustes Matching Algorithm. *Information Processing in Medical Imaging*, 1230:29–42, 1997.

- [94] Alexis Roche, Xavier Pennec, Grgoire Malandain, and Nicholas Ayache. Rigid registration of 3d ultrasound with mr images: a new approach combining intensity and gradient information. *IEEE Transactions on Medical Imaging*, 20:1038–1049, 2001.
- [95] Torsten Rohlfing. Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable. *IEEE Trans. Med. Imaging*, 31(2):153–163, 2012.
- [96] Daniel Rueckert, L. I. Sonoda, Carmel Hayes, Derek L. G. Hill, Martin O. Leach, and David J. Hawkes. Non-rigid registration using free-form deformations: Application to breast mr images. *IEEE Trans. Med. Imaging*, 18(8):712–721, 1999.
- [97] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the ICP algorithm. *Proceedings Third International Conference on 3D Digital Imaging and Modeling*, pages:145–152, 2001.
- [98] R B Rusu, N Blodow, and M Beetz. Fast Point Feature Histograms (FPFH) for 3D registration. *Proceedings of the IEEE International Conference on Robotics and Automation (2009)*, pages 3212–3217, 2009.
- [99] P H Schönemann. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31(1):1–10, 1966.
- [100] G L Scott and H C Longuet-Higgins. An algorithm for associating the features of two images. *Proceedings of the Royal Society B: Biological Sciences*, 244(1309):21–26, 1991.
- [101] Michaël Sdika. A fast nonrigid image registration with constraints on the jacobian using large scale constrained optimization. *IEEE Trans. Med. Imaging*, 27(2):271–281, 2008.
- [102] Thomas W. Sederberg and Scott R. Parry. Free-form deformation of solid geometric models. In *SIGGRAPH*, pages 151–160, 1986.
- [103] A. Segal, D. Haehnel, and S. Thrun. Generalized-icp. In *Proceedings of Robotics: Science and Systems*, Seattle, USA, June 2009.

- [104] C E Shannon. *The Mathematical Theory of Communication*, volume 27. University of Illinois Press, 1949.
- [105] L S Shapiro and J M Brady. Feature-based correspondence: an eigenvector approach. *Image and Vision Computing*, 10(5):283–288, 1992.
- [106] G.C. Sharp, S.W. Lee, and D.K. Wehe. Icp registration using invariant features. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(1):90 – 102, jan 2002.
- [107] Jianbo Shi and Carlo Tomasi. Good features to track. In *1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pages 593 – 600, 1994.
- [108] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, 2013.
- [109] C Studholme, D L Hill, and D J Hawkes. Automated 3-D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.
- [110] C Studholme, D L Hill, and D J Hawkes. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Medical Physics*, 24(1):25–35, 1997.
- [111] C Studholme, D L G Hill, and D J Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32(1):71–86, 1999.
- [112] Colin Studholme, Corina S. Drapaca, Bistra Iordanova, and Valerie Cardenas. Deformation-based mapping of volume change from serial brain mri in the presence of local tissue contrast change. *IEEE Trans. Med. Imaging*, 25(5):626–639, 2006.
- [113] Richard Szeliski. Image alignment and stitching: A tutorial. *Found. Trends. Comput. Graph. Vis.*, 2(1):1–104, January 2006.
- [114] A Tekbas, A Seitel, A M Franz, Boris Radeleff, F Pianka, and S A Müller. Computer-Assisted Needle Insertion for Abdominal Interventions. *IFMBE Proceedings*, (Vol. 25):159–162, 2009.

- [115] Jean-Philippe Thirion. Non-rigid matching using demons. In *CVPR*, pages 245–251, 1996.
- [116] Jean-Philippe Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.
- [117] Michael Unser. Splines: A perfect fit for signal/image processing. *IEEE Signal Processing Magazine*, 16:22–38, 1999.
- [118] P A van den Elsen, E J. D Pol, and M A Viergever. Medical image matching-a review with classification. *IEEE International Conference of the Engineering in Medicine and Biology Society*, 12(1):26–39, 1993.
- [119] P Viola and W M Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):16–23, 1997.
- [120] Frank K Wacker, Sebastian Vogt, Ali Khamene, John A Jesberger, Frank Sauer, Jeffrey L Duerk, and Jonathan S Lewin. An Augmented Reality System for MR Image guided Needle Biopsy : Initial Results in a Swine Methods : Results : Conclusion. *Radiology*, 238(2):497–504, 2006.
- [121] G Wahba. *Spline Models for Observational Data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, 1990.
- [122] Michael J Wallace, Sanjay Gupta, and Marshall E Hicks. Out-of-plane computed-tomography-guided biopsy using a magnetic-field-based navigation system. *CardioVascular and Interventional Radiology*, 29(1):108–113, 2006.
- [123] Wolfgang Wein, Shelby Brunke, Ali Khamene, Matthew R. Callstrom, and Nasir Navab. Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical Image Analysis*, (5):577–585.
- [124] W M Wells, P Viola, H Atsumi, S Nakajima, and R Kikinis. Multi-modal volume registration by maximization of mutual information., 1996.
- [125] R T Whitaker, C Crampton, D E Breen, M Tuceryan, and E Rose. Object Calibration for Augmented Reality. *Computer Graphics Forum*, 14(3):15–28, 1995.

- [126] Andrew D. Wiles, A. Likholyot, Don D. Frantz, and Terry M. Peters. A statistical model for point-based target registration error with anisotropic ducial localizer error. *IEEE Trans. Med. Imaging*, 27(3):378–390, 2008.
- [127] R P Woods, S R Cherry, and J C Mazziotta. Rapid automated algorithm for aligning and reslicing PET images. *Journal Of Computer Assisted Tomography*, 16(4):620–633, 1992.
- [128] R P Woods, J C Mazziotta, and S R Cherry. MRI-PET registration with automated algorithm. *Journal Of Computer Assisted Tomography*, 17(4):536–546, 1993.
- [129] Lyubomir Zagorchev and A. Ardeshir Goshtasby. A comparative study of transformation functions for nonrigid image registration. *IEEE Transactions on Image Processing*, 15(3):529–538, 2006.
- [130] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.
- [131] Xiahai Zhuang, Simon R. Arridge, David J. Hawkes, and Sébastien Ourselin. A nonrigid registration framework using spatially encoded mutual information and free-form deformations. *IEEE Trans. Med. Imaging*, 30(10):1819–1828, 2011.
- [132] B Zitova and Flusser J. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.

Appendix A

Image registration results

We report here the results of the registration algorithm that we have described in chapter 2. The tests were performed using MATLAB environment and FAIR toolbox [78].

A.1 Parametric registration using mutual information

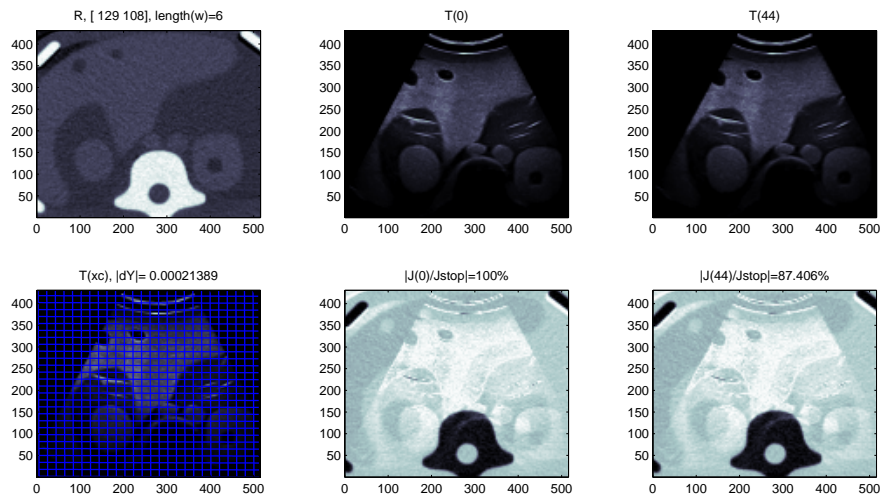


Figure A.1: Af ne registration using MI. Top row, from left t o right: the reference image, the target image before the rst iteration, the target t image after 44 iterations. Bottom row, from left to right: the target image after the registration with the deformation grid superimposed, target and reference difference image before registration, target and reference difference image after registration.

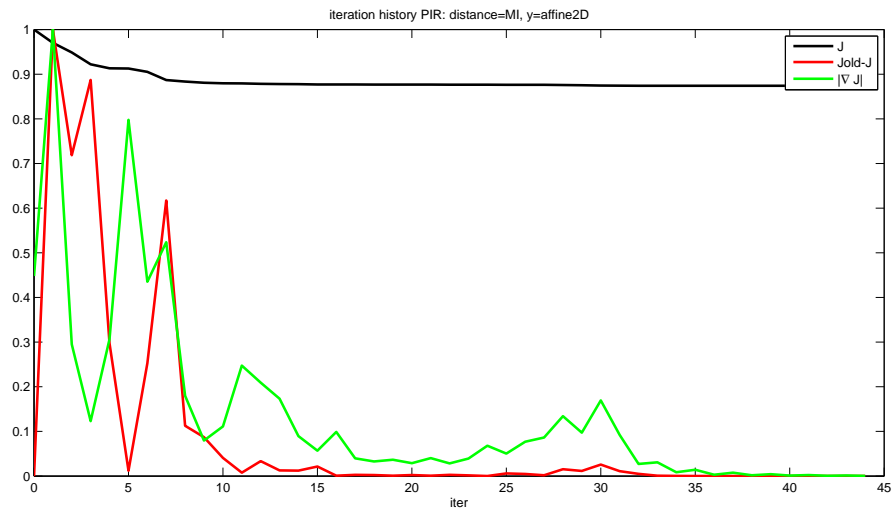


Figure A.2: Af ne registration using MI. The graphical trend of the objective function J during 44 iterations (the black line). The green line gives the graph of $|\nabla J|$, while the red shows line the graph of the difference between the previous J value and the current value.

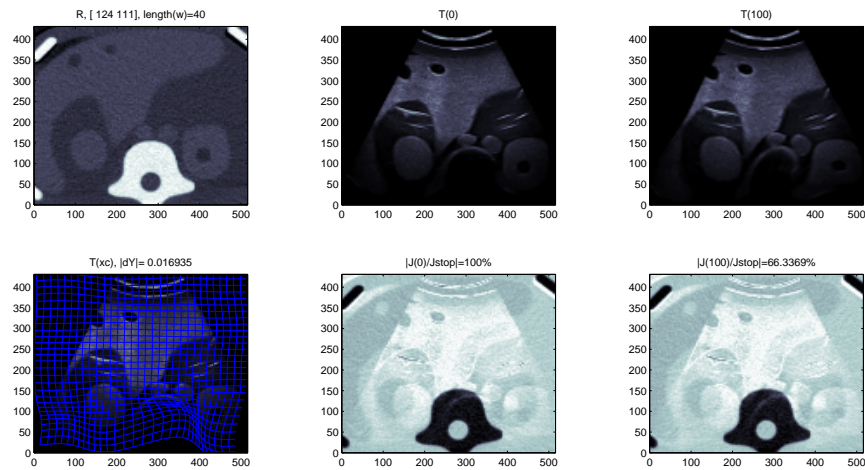


Figure A.3: Spline registration using MI. Top row, from left to right: the reference image, the target image before the first iteration, the target image after 100 iterations. Bottom row, from left to right: the target image after the registration with the deformation grid superimposed, target and reference difference image before registration, target and reference difference image after registration.

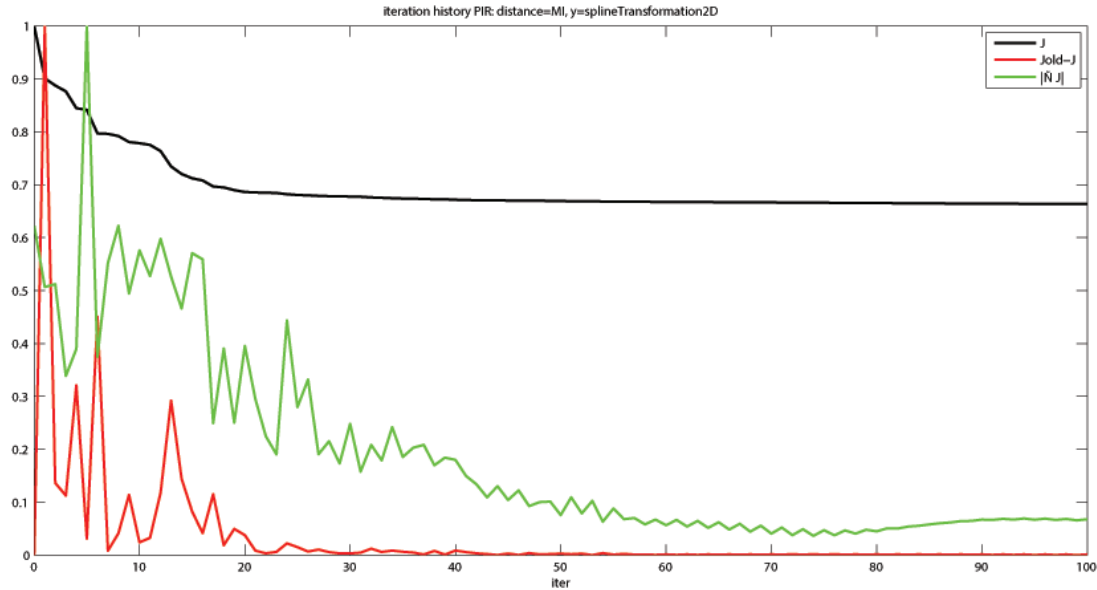


Figure A.4: Spline registration using MI. The graphical trend of the objective function J during 100 iterations (the black line). The green line gives the graph of $|\nabla J|$, while the red shows line the graph of the difference between the previous J value and the current value.

The results obtained using MI as distance measure are not satisfactory in both affine and spline cases (figures A.1 and A.3). The optimization process tends very quickly to a local minimum, and the transformation leaves the image almost unchanged, in the case of affine transformation (figure A.1) or applies a very small local deformation, in the case of the spline deformation (figure A.3).

The iterations can stop when the error is less than a threshold, as in the affine case, or, in the case that this threshold is not reached, we imposed a limit of 100 iterations.

The trend of the MI measure is to quickly reach a stable position (figures A.2 and A.4) at around 66% in the affine case, and 87% in the spline case, of the of the initial value.

The computation time is very long.

The minimization scheme used was Gauss-Newton.

A.2 Parametric registration using normalized gradient

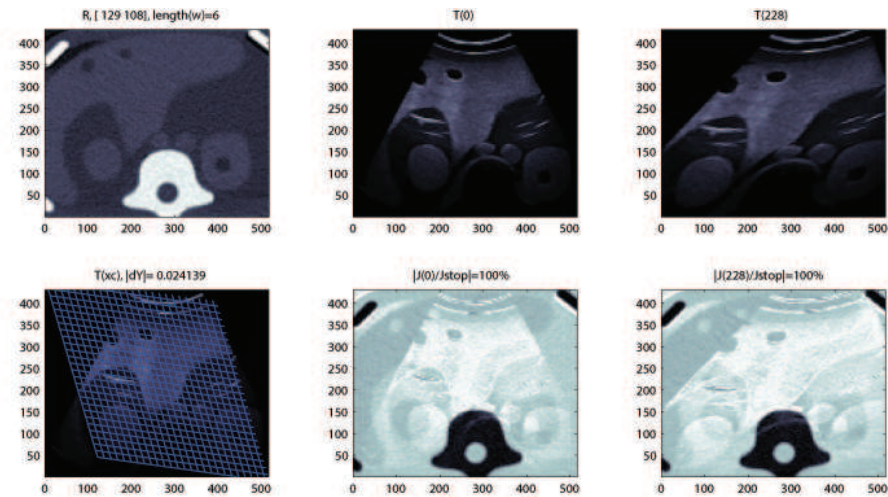


Figure A.5: Affine registration using NGF. The results are obtained after 228 iterations. The scheme is the same as in figure A.2

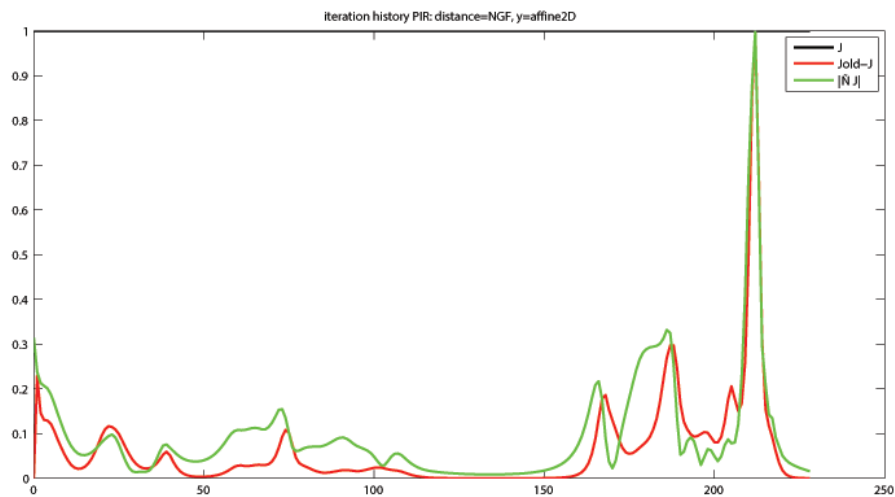


Figure A.6: Affine registration using NGF. The graphical trend of the objective function J during 228 iterations (the black line). The green line gives the graph of $|\nabla J|$, while the red line shows the graph of the difference between the previous J value and the current value.

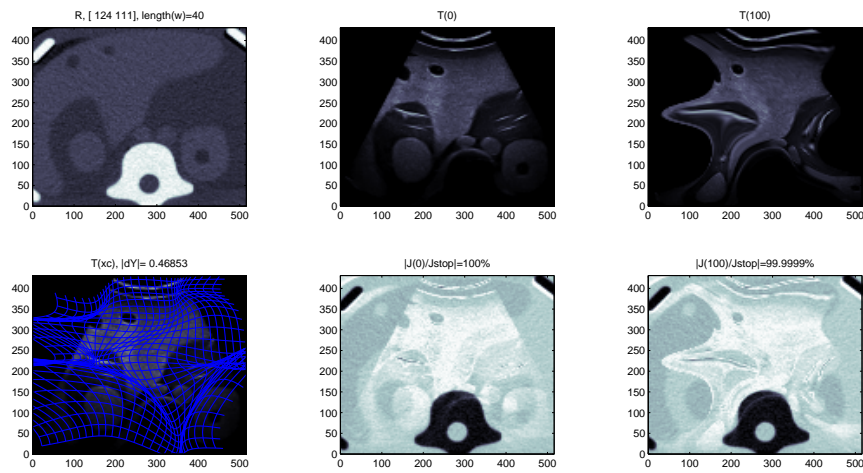


Figure A.7: Spline registration using NGF. Top row, from left to right: the reference image, the target image before the first iteration, the target image after 100 iterations. Bottom row, from left to right: the target image after the registration with the deformation grid superimposed, target and reference difference image before registration, target and reference difference image after registration.

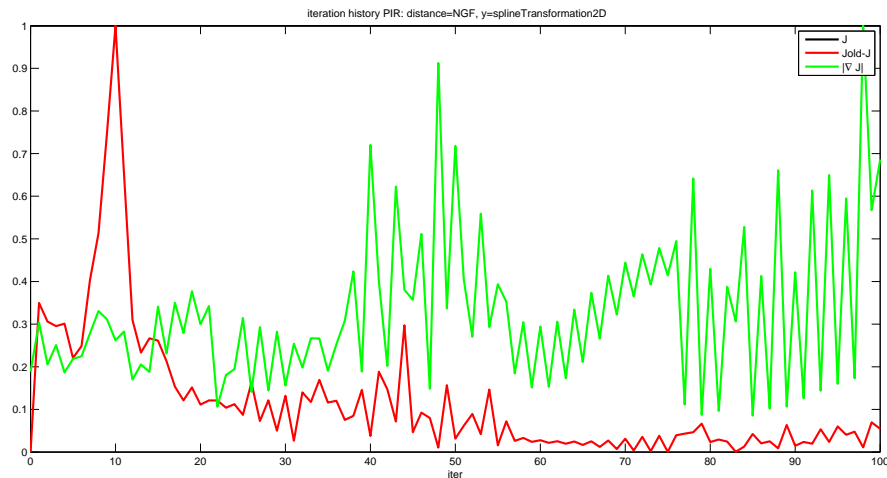


Figure A.8: Spline registration using NGF. The graphical trend of the objective function J during 100 iterations (the black line). The green line gives the graph of $|\nabla J|$, while the red shows line the graph of the difference between the previous J value and the current value.

After our tests, it turned out that the NGF distance measure is not a good choice, at least for the images we have chosen. As it can be seen in figures A.6 and A.8, the values of the objective function do not change over the iterations, they are indistinguishable in the graphic. In fact, they remain always very near to the initial value 1. Visually, the figures A.5 and A.7 show that in both cases the results are meaningless and the deformation applied is arbitrary.

The only advantage over the previous approach was the computation time.

Also in this case we have used Gauss-Newton for the minimization.

A.3 Multilevel parametric and non parametric registration

A.3.1 Input

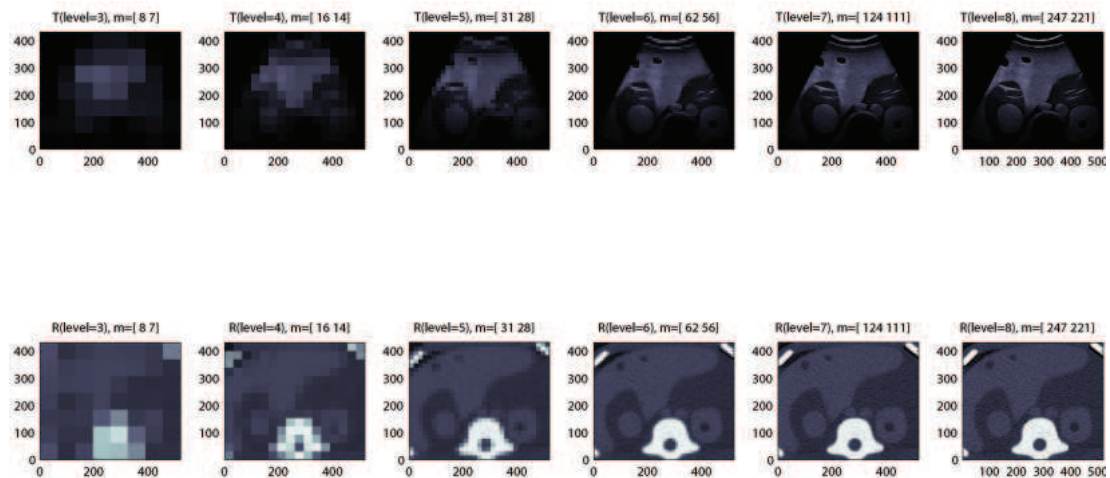


Figure A.9: The target (top row) and the reference (bottom row) images represented using 6 different levels.

The multilevel representation of the input data is required first of all in order to reduce the risk of being trapped by local minimum. The target and the reference images are depicted in the figure A.9. At the same time, a solution of a coarse representation of the problem serves as a starting point for a representation with more details. Starting with a very coarse representation, the procedure is repeated on each level, until all details provided by the initial data are resolved.

From an optimization point of view, the multilevel representation yield a smoother representation of the objective function. A smooth problem may be easier to resolve and, based on a good starting point, the more detailed problem can be solved quicker.

At the same time, the multilevel representation is useful to reduce the computational time. The complexity of the registration algorithms increase exponentially with the dimension of the input data.

A.3.2 Parametric multilevel spline registration

Figure ?? reports the results of the multilevel spline registration using the MI distance measure. After 10 iteration on each level, the result is passed to the first iteration of the next level.

The transformations also in these cases are not natural, after a strong bending obtained at the first level, the other levels tend to bend less than the image.

The computation time is very high and an ulterior refinement at each level through the other iterations does not improve the results.

The tests using the affine transformation are not reported here but the results are similar to the non-level approach, the transformation blocks quickly into a local minimum.

The tests done using NGF distance measure give again arbitrary deformations.

A.3.3 Non parametric multilevel registration

Elastic

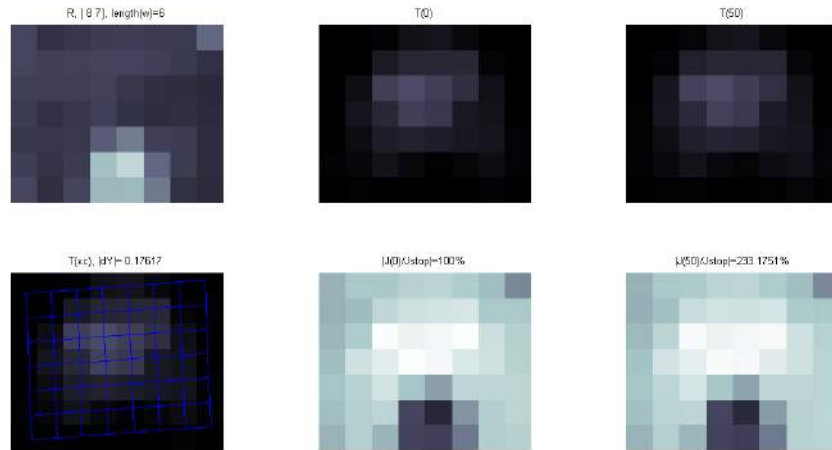


Figure A.10: The initialization of the algorithm at the first level considered uses the affine transformation.

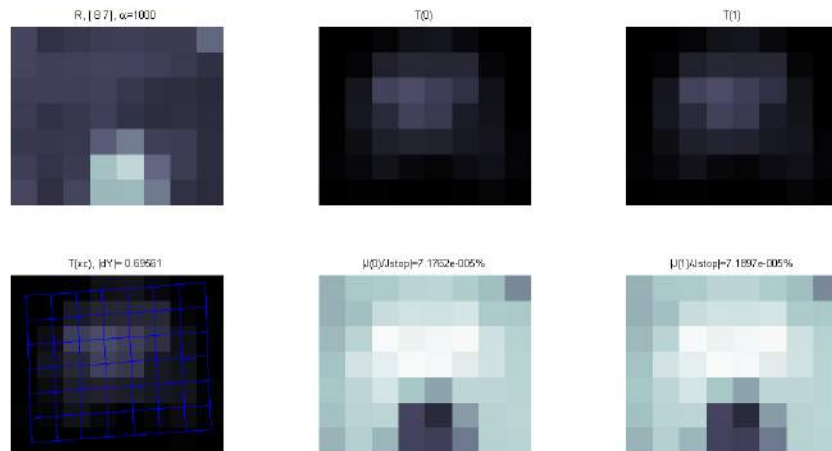


Figure A.11: First level of registration using the affine transformation as initialization and elastic registration to refine.

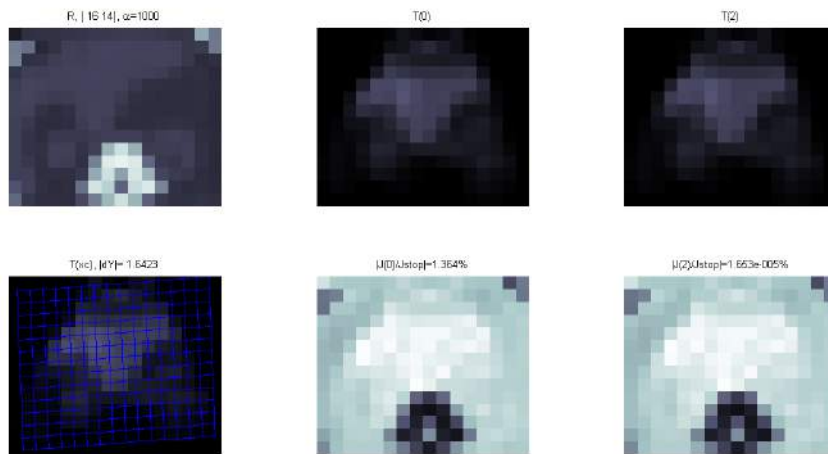


Figure A.12: Second level of registration: the initialization uses the previous obtained transformation.

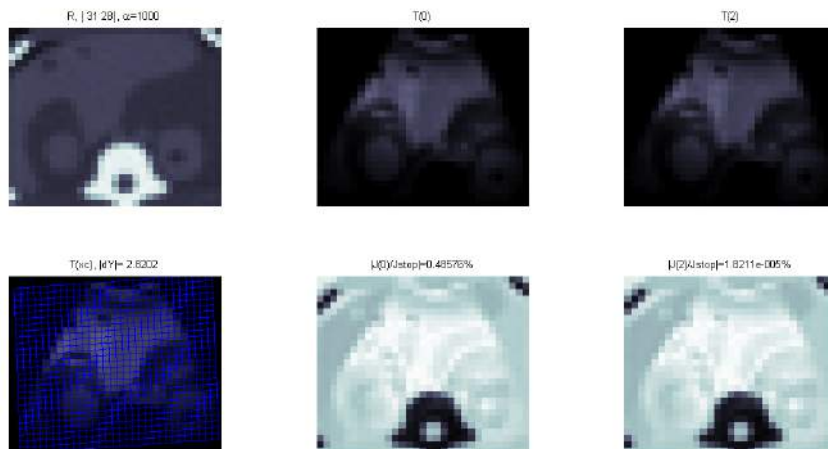


Figure A.13: Third level of registration: the initialization uses the previous obtained transformation.

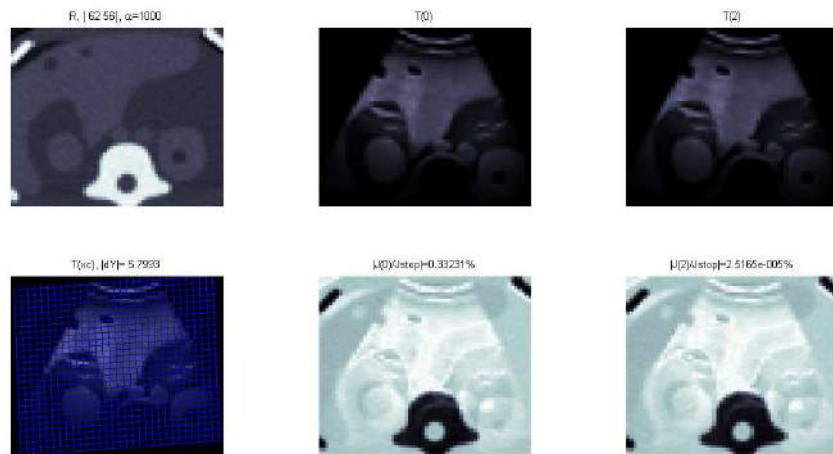


Figure A.14: Fourth level of registration: the initialization uses the previous obtained transformation.

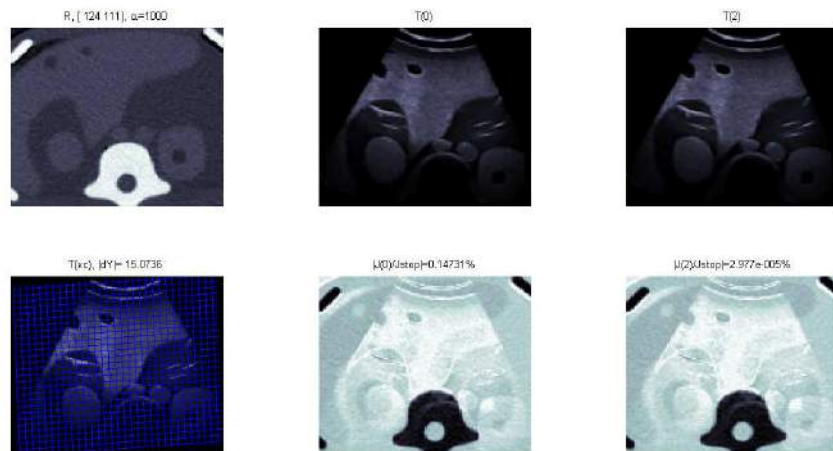


Figure A.15: Last level of registration: the initialization uses the previous obtained transformation.

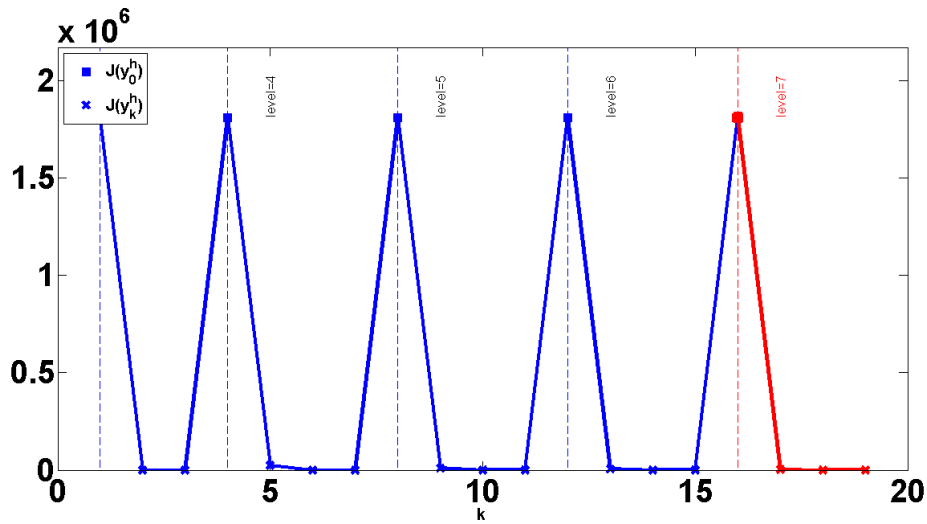


Figure A.16: Iteration history of multilevel non-parametric elastic registration: vertical lines separate different levels. The squares represent the initial value of MI, while the crosses show the value on each iteration.

Figures A.10 to A.16 show the results of the non-parametric elastic multi-level registration. The algorithm starts at the coarser level using an affine transformation, then the registration at each level starts using the previous obtained result. After the initialization, there is not much interaction of the algorithm with the data, the images remain almost unchanged and the number of iterations is very low on each level. Even so, the computation time is very high.

Curvature

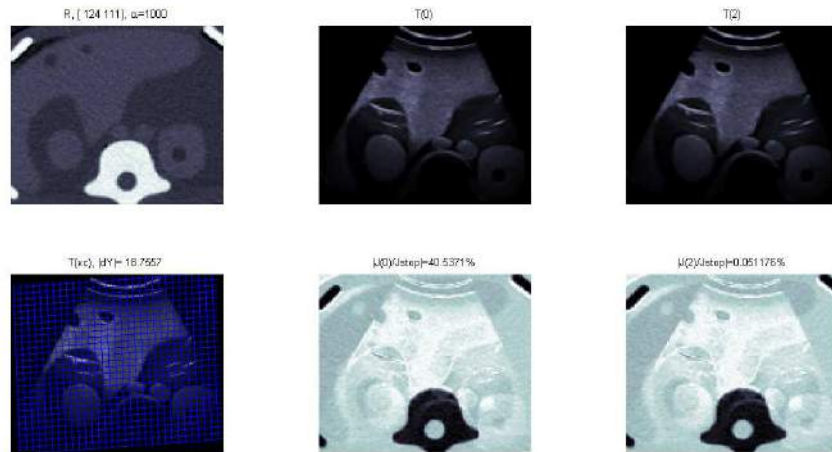


Figure A.17: The result of the registration using the curvature.

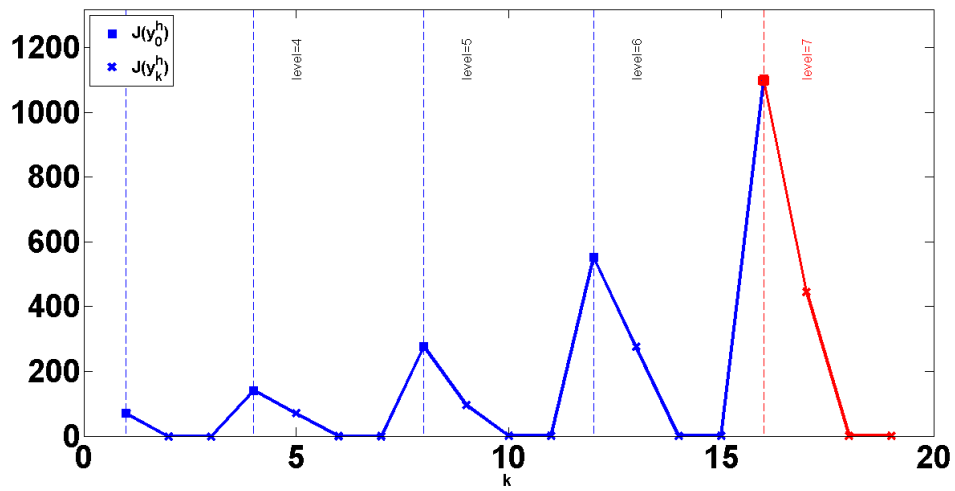


Figure A.18: Iteration history of multilevel non-parametric curvature registration: vertical lines separate different levels. The squares represents the initial value of MI, while the crosses show the value on each iteration.

The results obtained using the curvature registration are very similar with those obtained using the elastic registration therefore we report here only the registration on the last level (gure A.17) and the graphic of the iterations (gure A.18).

A.4 Computation time

Even if in the implementation of the registration algorithms we used Matlab software, we report in the table A.1 the time taken for each of the test we have presented in the previous sections.

Registration type	Transformation type	Time (sec.)
Parametric MI	Affine	15
	Spline	87
Parametric NGF	Affine	6
	Spline	18
Multilevel parametric	Affine	29
	Spline	143
Multilevel non-parametric	Elastic	214
	Curvature	168

Table A.1: Computation time of the principal registration algorithm tested.

Appendix B

Landmark-based registration results

We report here the results of the registration algorithms that we have described in chapter 3.

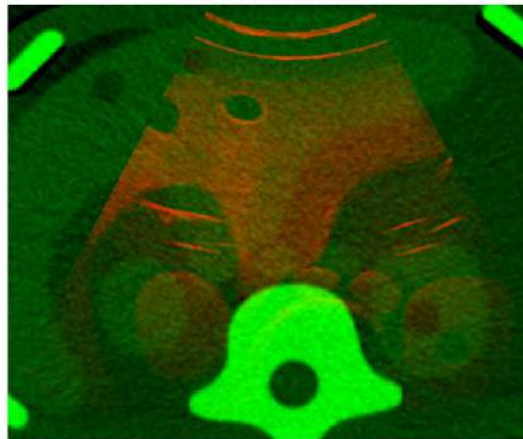


Figure B.1: Overlapped images before registration: target image (US) is represented by the red layer, source image (CT) by the green layer.

B.1 Landmark-based registration: linear

The linear registration in 2D involves the computation of 6 parameters, 2 for translation, 1 for rotation and 2 for scaling therefore a minimum number of 3 landmarks is required (Figure B.2).

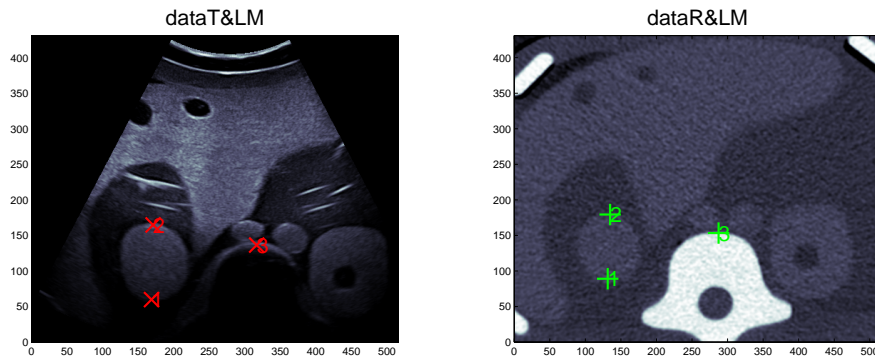


Figure B.2: The template and the reference images with 3 corresponding pair of landmarks.

Even if the solution with 3 landmarks is in closed form, the registration works well only in the region where the landmarks are placed (Figure B.3 and B.4, left). The closed form solution applied to the entire image is not satisfactory (B.4).

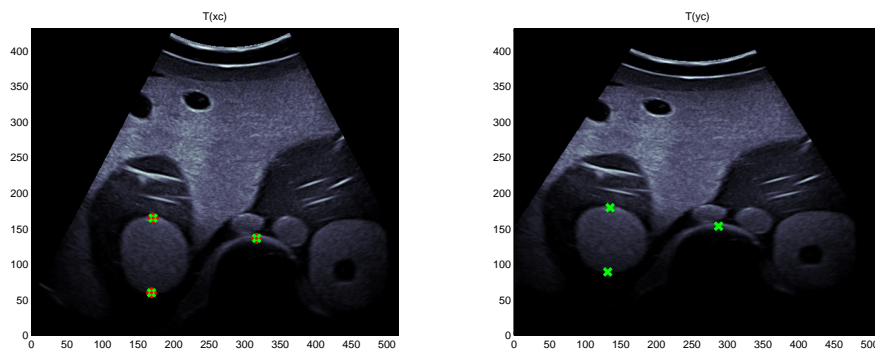


Figure B.3: The template image before and after the linear registration with the registered landmarks highlighted.

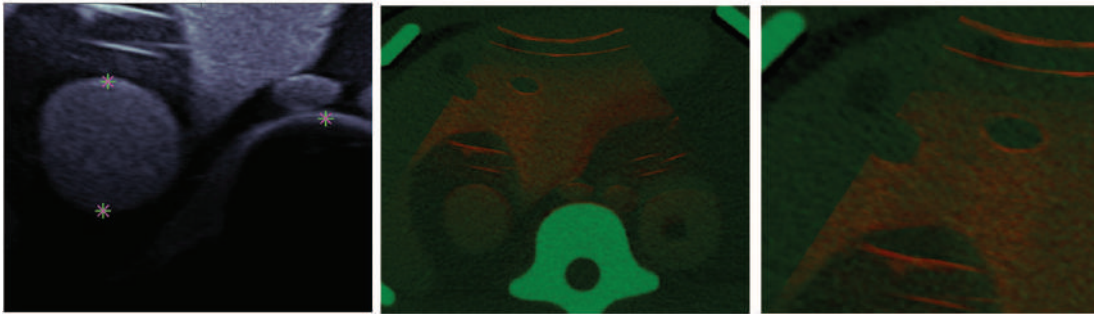


Figure B.4: From left to right: a close-up of the linearly registered landmarks, the target (red) and reference (green) images superimposed after the linear registration and a close-up of the same image.

Increasing the number of landmarks (Figure B.5), the registration problem becomes overdetermined and the solution is given by minimizing the sum of the distances between every pair of corresponding landmarks.

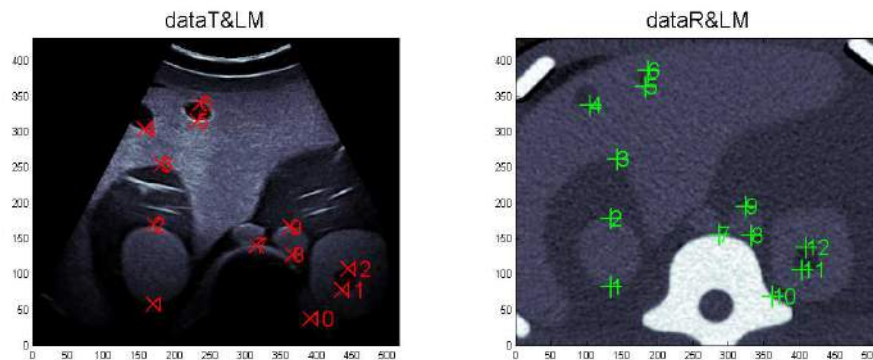


Figure B.5: The template and the reference images with 12 corresponding landmarks.

Even though not all the landmarks are aligned (Figure B.6 and B.7 left), the result improve a lot compared to the closed form solution (Figure B.7, middle and right).

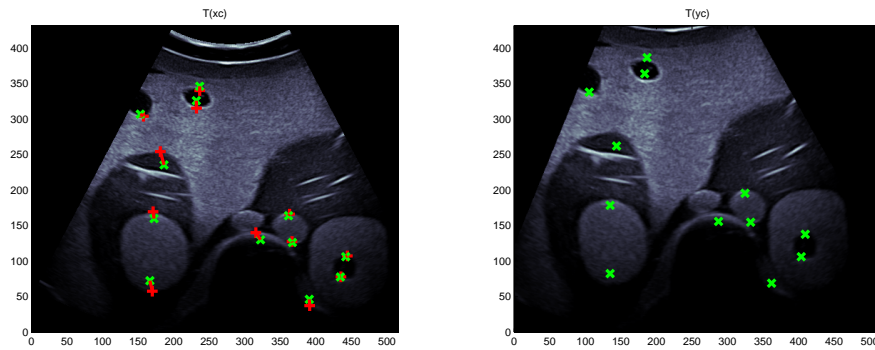


Figure B.6: The template image before and after the registration with the registered landmarks highlighted.

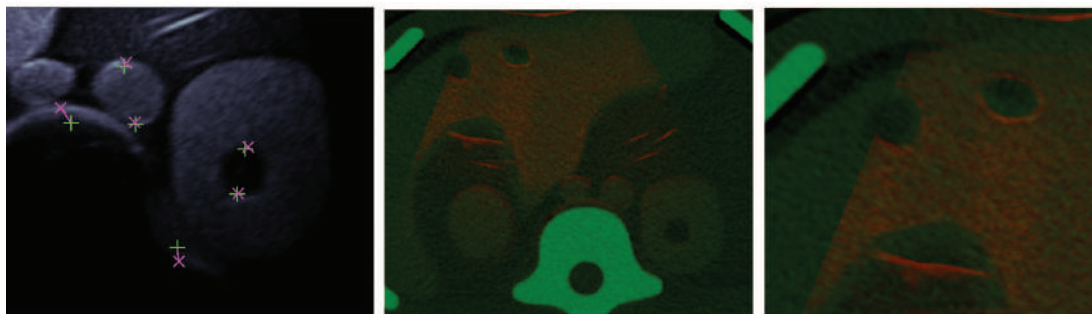


Figure B.7: From left to right: a close-up of the registered landmarks, the target and reference images superimposed after the registration, a close-up of the superimposed images.

B.2 Landmark-based registration: non-linear

B.2.1 Quadratic registration

The results of the linear registration may be improved by using a non-linear approach. The quadratic solution in 2D has a number of 12 parameters and is completely given by 6 landmarks (Figure B.8). In this case the error in the alignment of the landmarks is negligible because depends only on the computational precision (Figure B.8).

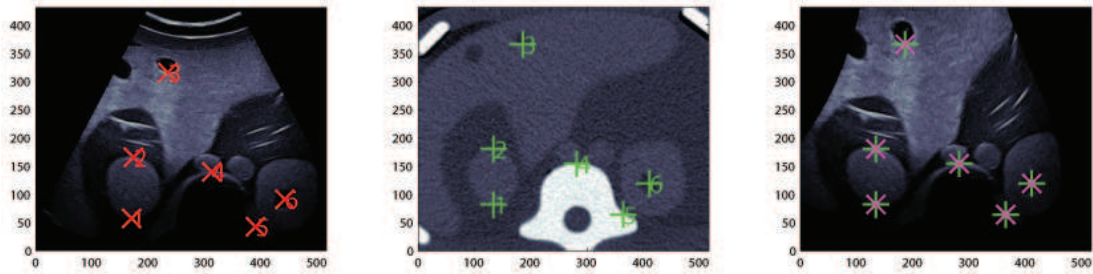


Figure B.8: The target and reference image with 6 selected landmarks. On the right the target image registered and the 6 landmarks perfectly aligned by the quadratic model.

The registration result applied to the entire image, even in the 6 landmarks case, is acceptable (Figure B.9).

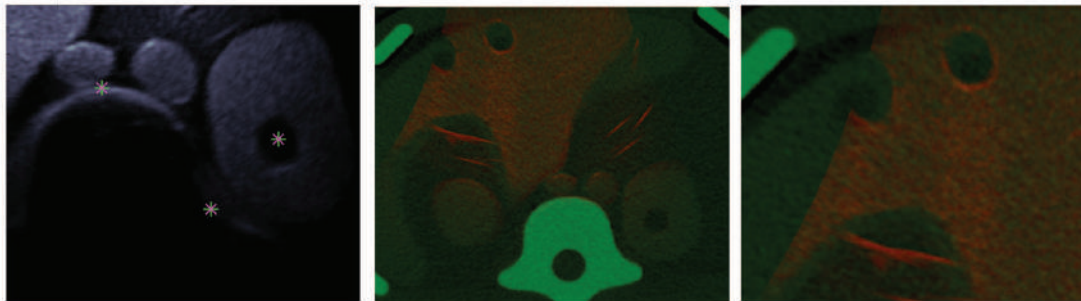


Figure B.9: The two images overlapped after the quadratic registration with 6 landmarks in the middle and a close-up on the right. On the left a close-up of some registered landmarks.

As in the linear case, increasing the number of landmarks, the solution is overdetermined but the overall result is better (Figure B.10). By using the quadratic term, the alignment of 15 landmarks gives a minor error compared with the use of 12 landmarks in the linear case (Figure B.11).

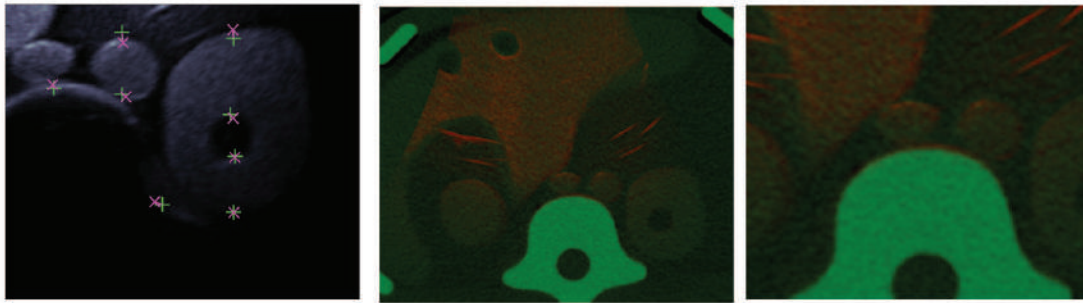


Figure B.10: The two images overlapped after the quadratic registration with 15 landmarks in the middle and a close-up on the right. On the left a close-up of some registered landmarks.

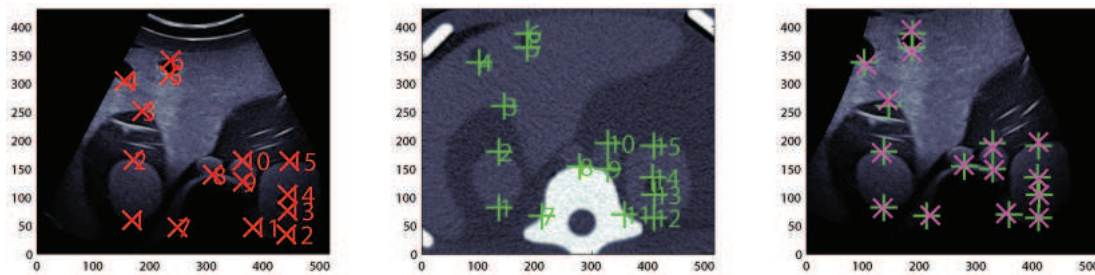


Figure B.11: The target and reference image with 15 selected landmarks. On the right the target image registered and the 15 landmarks aligned by the quadratic model.

B.2.2 Thin-plate-spline registration

The polynomial solution may produce arbitrary deformations especially when the number of landmarks is low.

The approach based on the thin-plate-spline (TPS) functions further refines the previous results. The spline functions may yield the perfect matching of the landmarks (interpolation) or an approximation, when the landmarks are not completely overlapped. These two conditions are realized by varying a smoothing parameter α . The solutions range from a low degree of bending when α is large to a high degree of bending in the case of small α . However, considering a number of landmarks from 6 to 15 and α ranging from 0 to 1000, this phenomenon is barely visible and the distance between the two sets of registered landmarks increases visibly only in the case of $\alpha = 1000$.

The figures B.12 to B.29 report the results for different values of α and different number of landmarks.

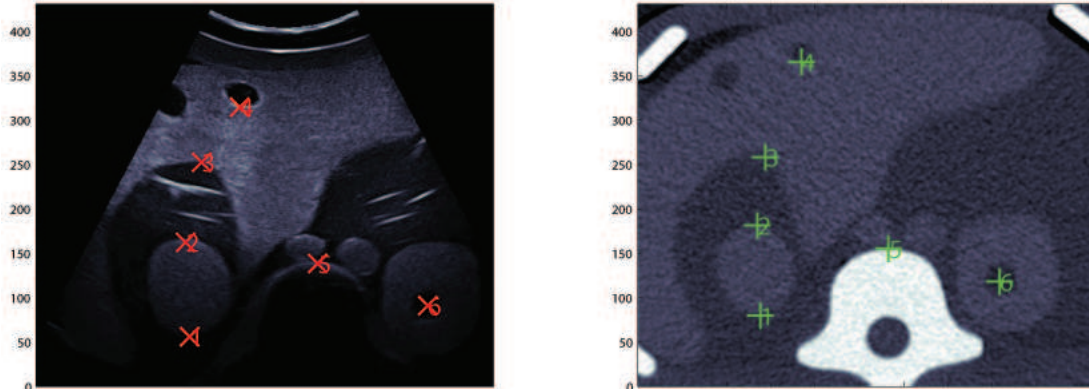


Figure B.12: The target (left) and reference (right) image with 6 selected landmarks.

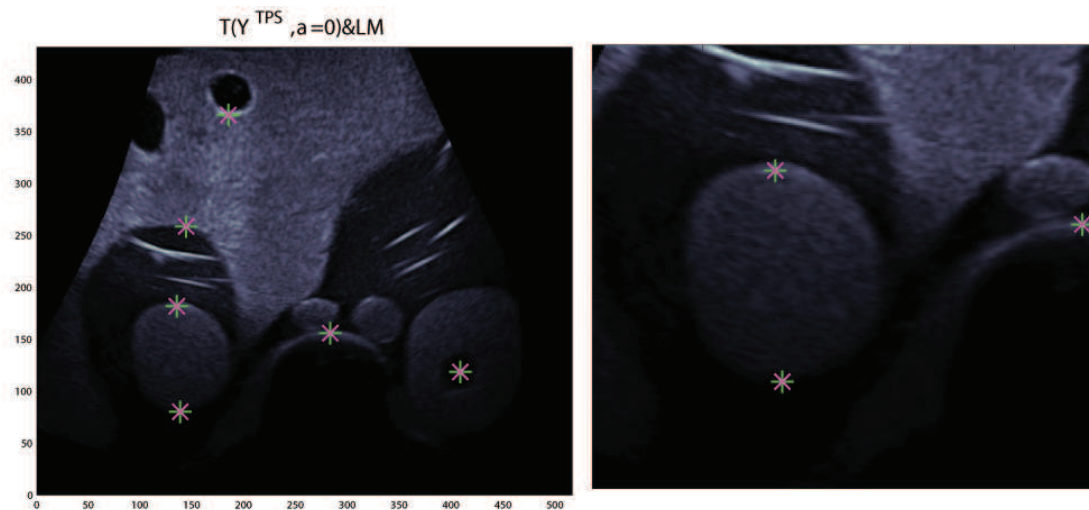


Figure B.13: The target image from g.B.12 registered for $\alpha = 0$

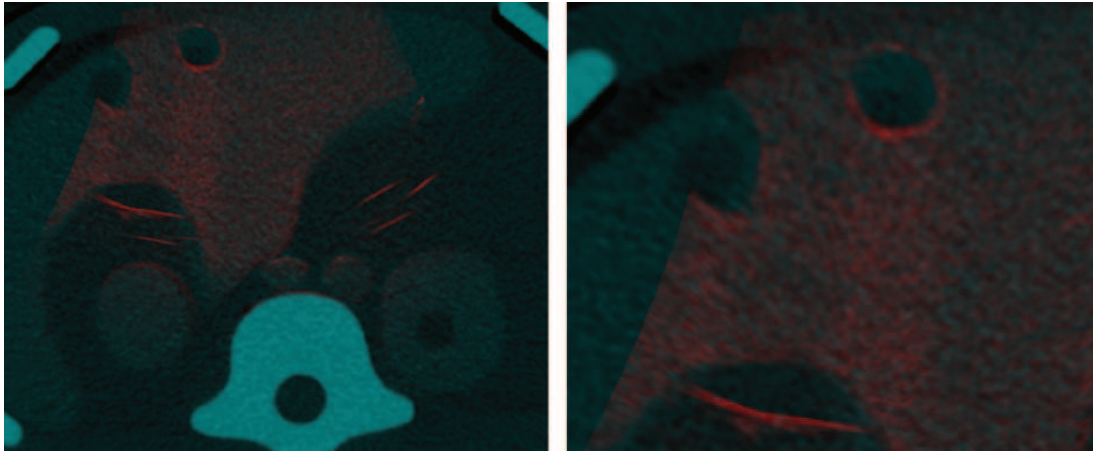


Figure B.14: The two images from g.B.12 overlapped after the T PS registration for $\alpha = 0$

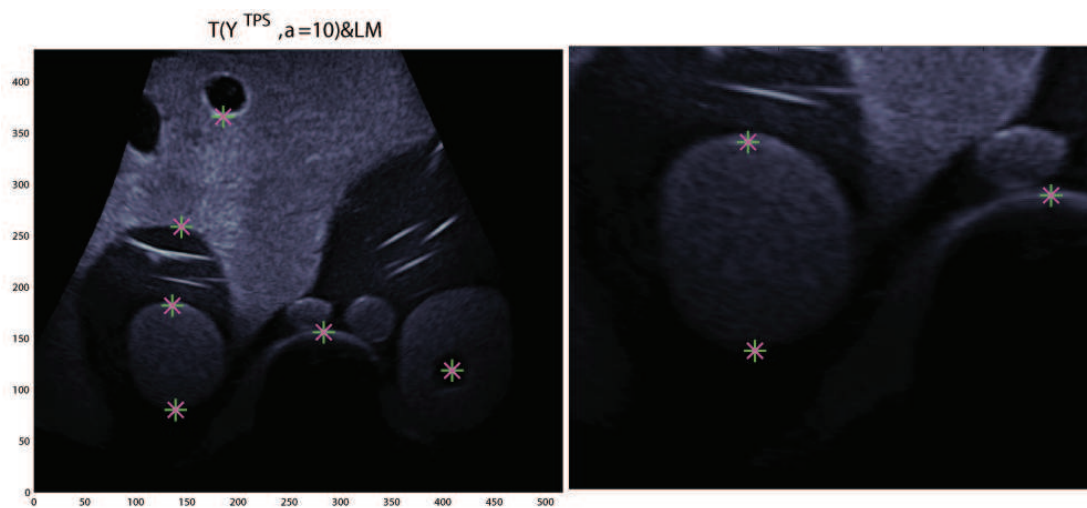


Figure B.15: The target image from g.B.12 registered for $\alpha = 10$

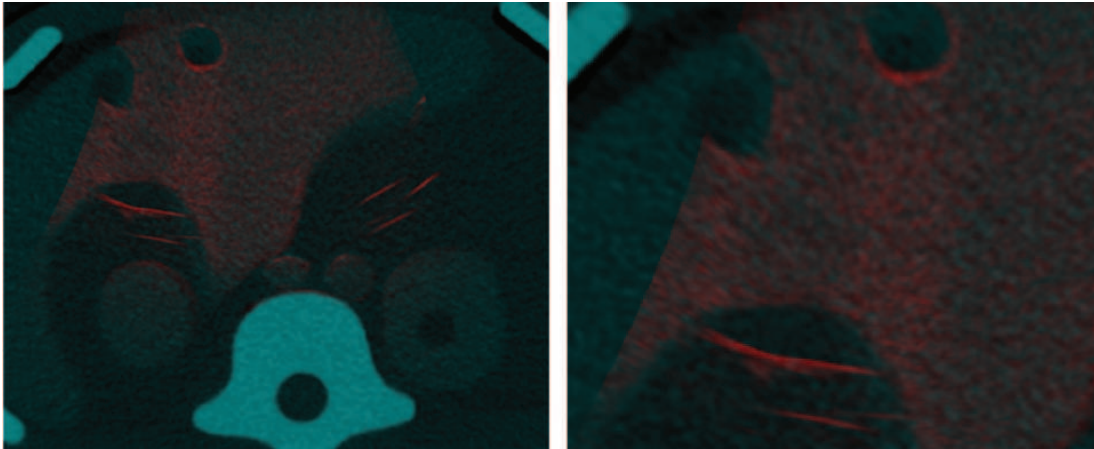


Figure B.16: The two images from g.B.12 overlapped after the T PS registration for $\alpha = 10$

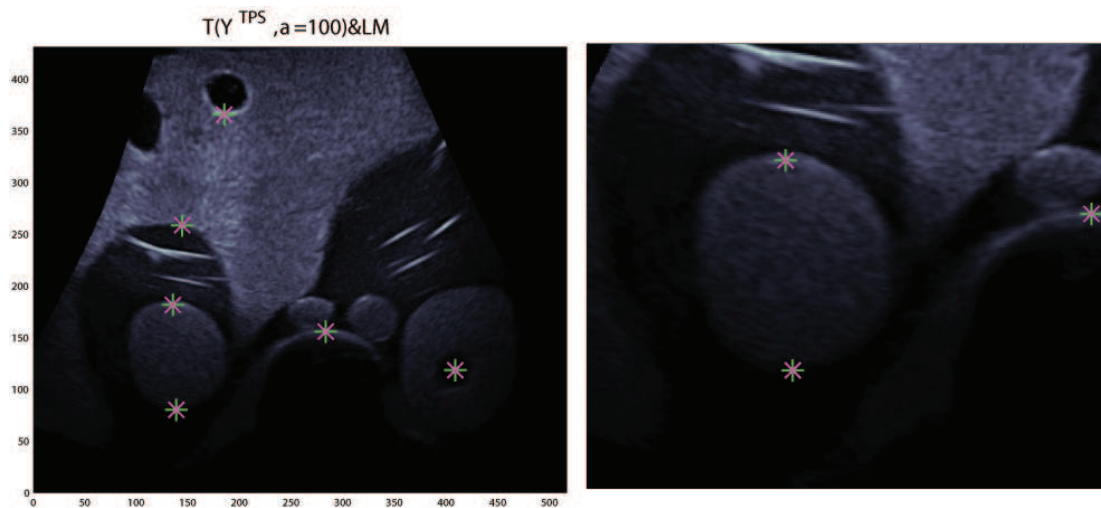


Figure B.17: The target image from g.B.12 registered for $\alpha = 100$

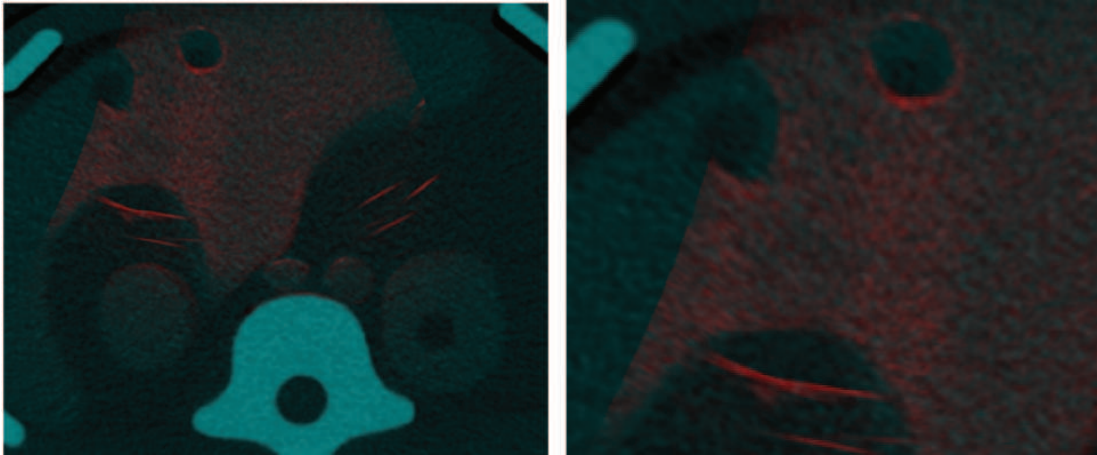


Figure B.18: The two images from g.B.12 overlapped after the T PS registration for $\alpha = 100$

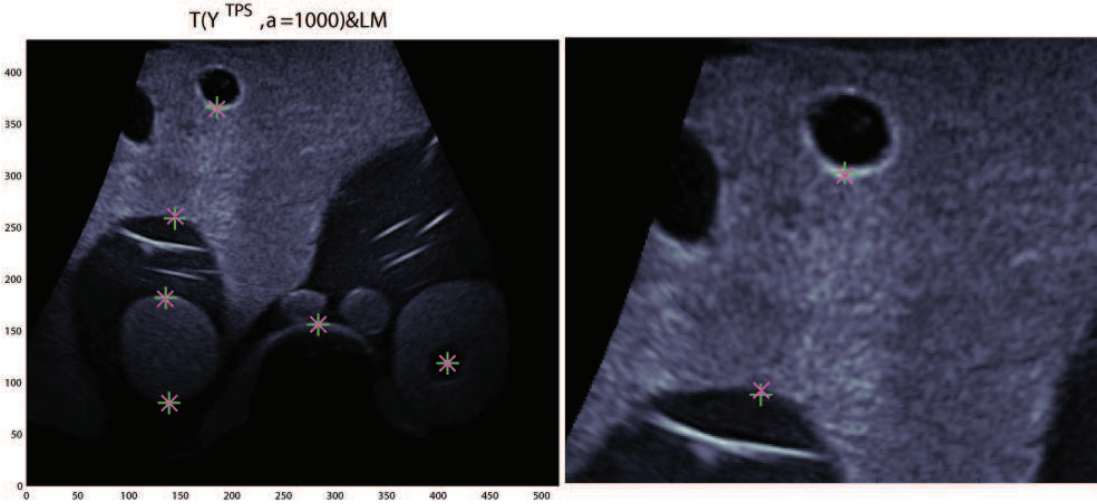


Figure B.19: The target image from g.B.12 registered for $\alpha = 1000$

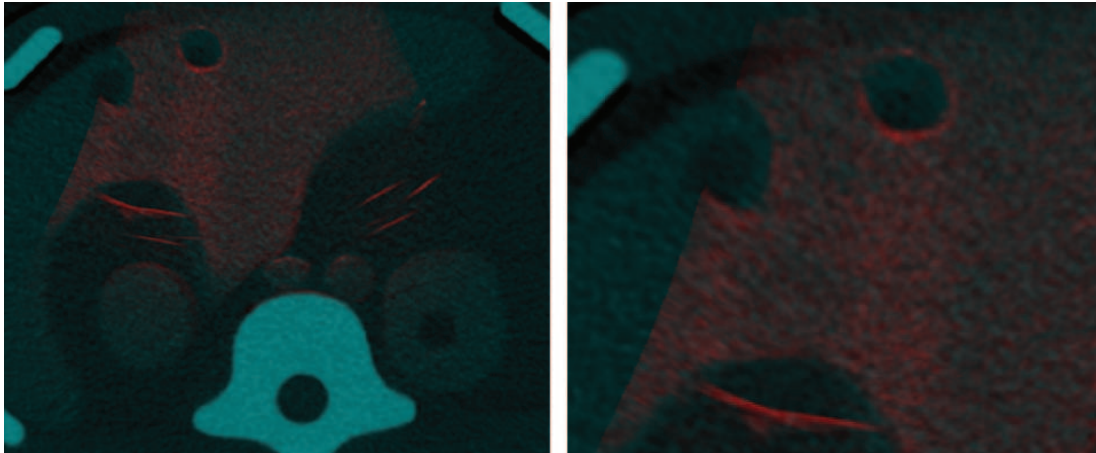


Figure B.20: The two images from g.B.12 overlapped after the T PS registration for $\alpha = 1000$

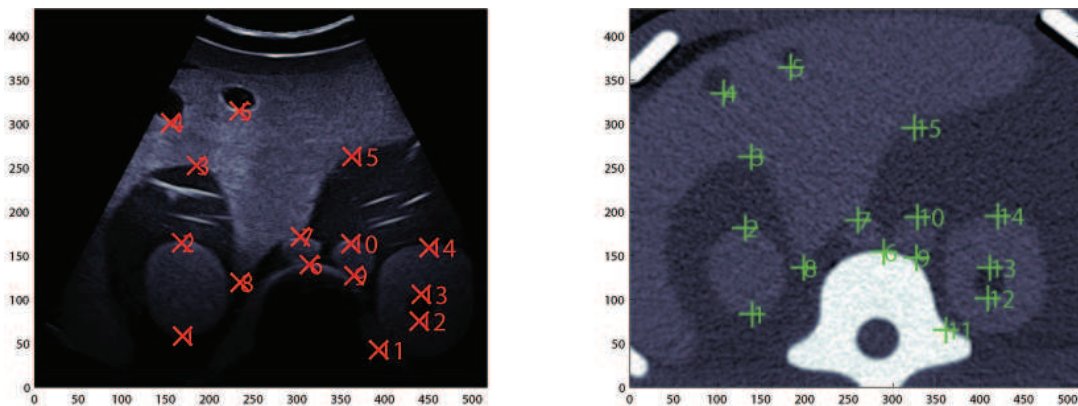


Figure B.21: The target (left) and reference (right) image with 15 selected landmarks.

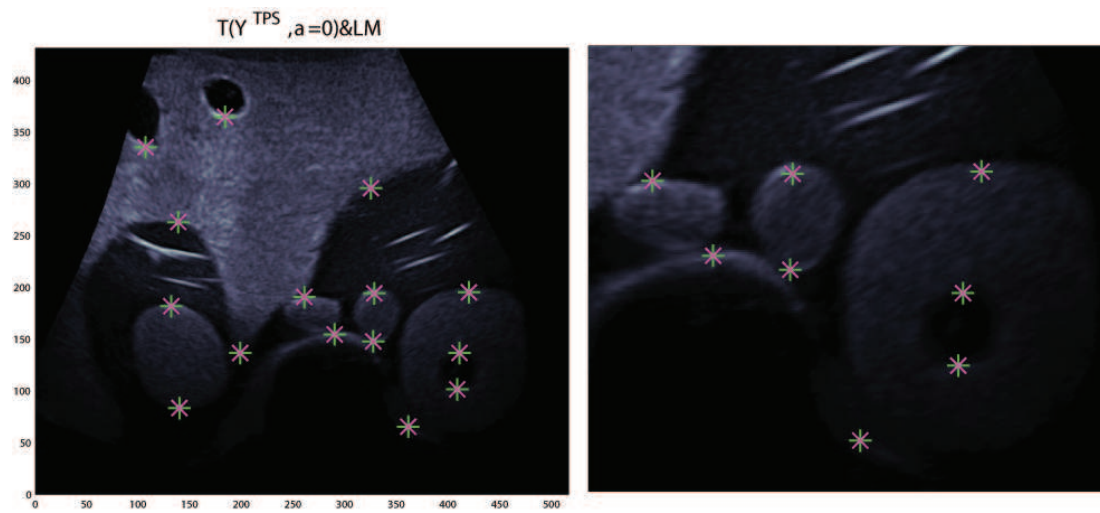


Figure B.22: The target image from g.B.21 registered for $\alpha = 0$

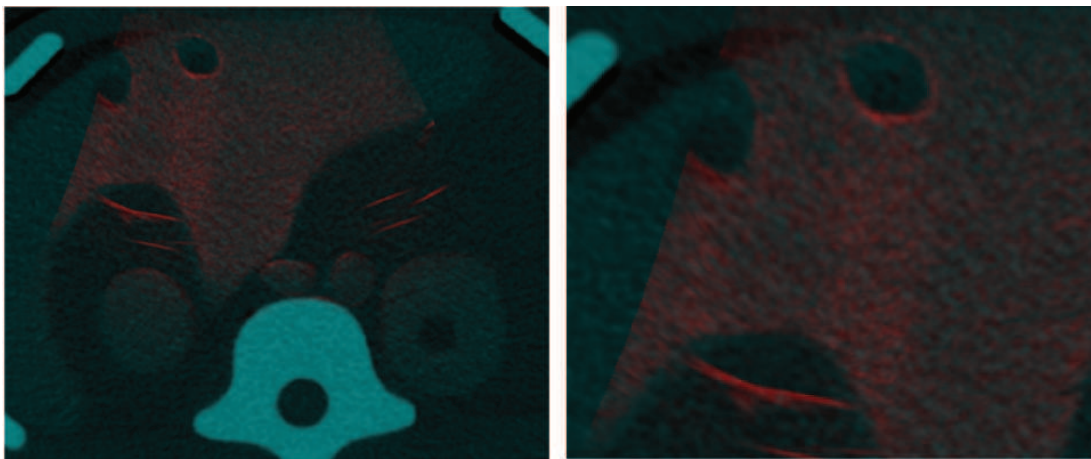


Figure B.23: The two images from g.B.21 overlapped after the TPS registration for $\alpha = 0$

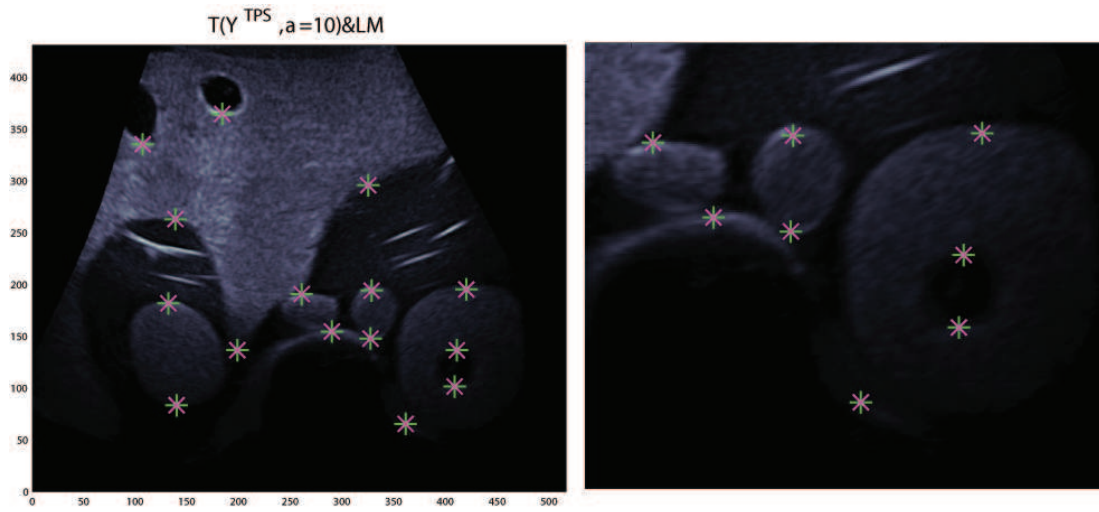


Figure B.24: The target image from g.B.21 registered for $\alpha = 10$

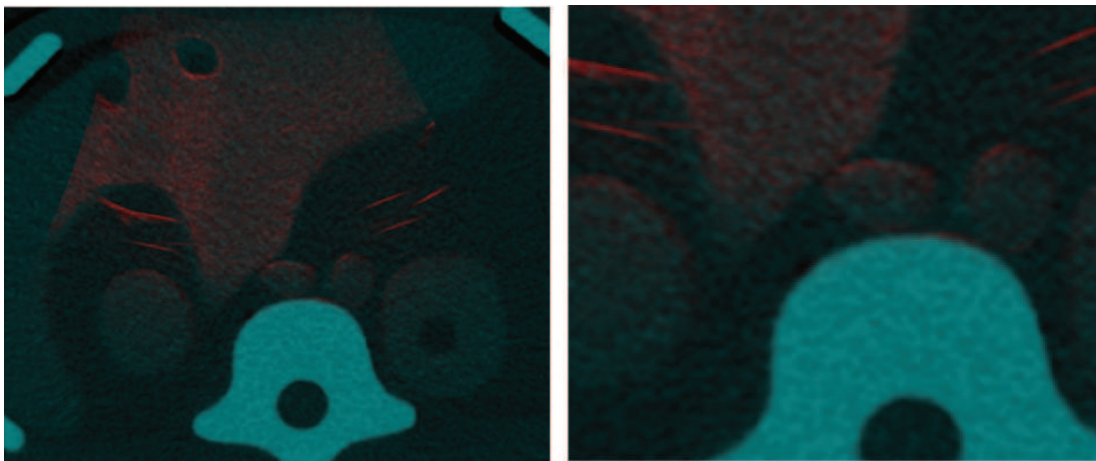


Figure B.25: The two images from g.B.21 overlapped after the T PS registration for $\alpha = 100$

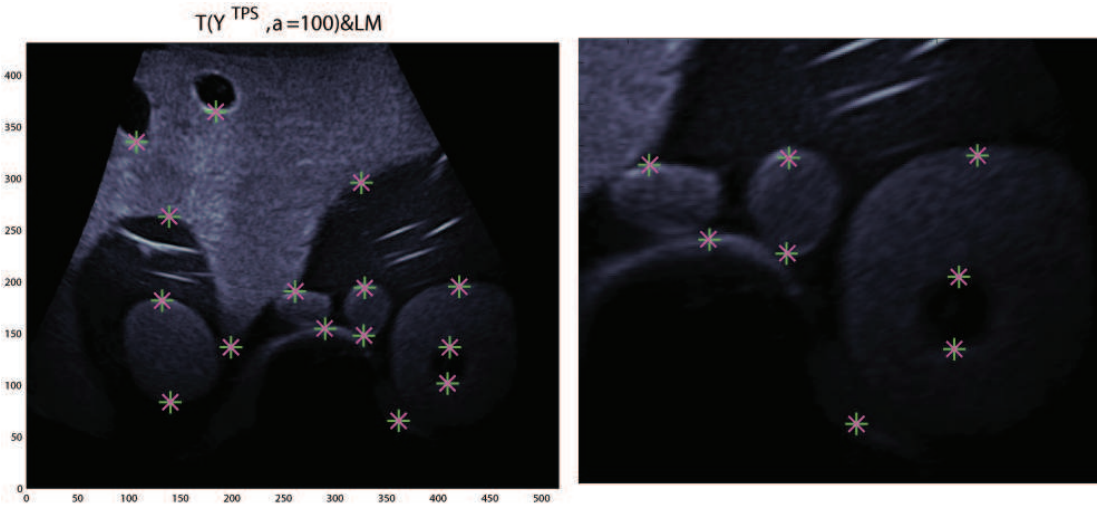


Figure B.26: The target image from g.B.21 registered for $\alpha = 100$

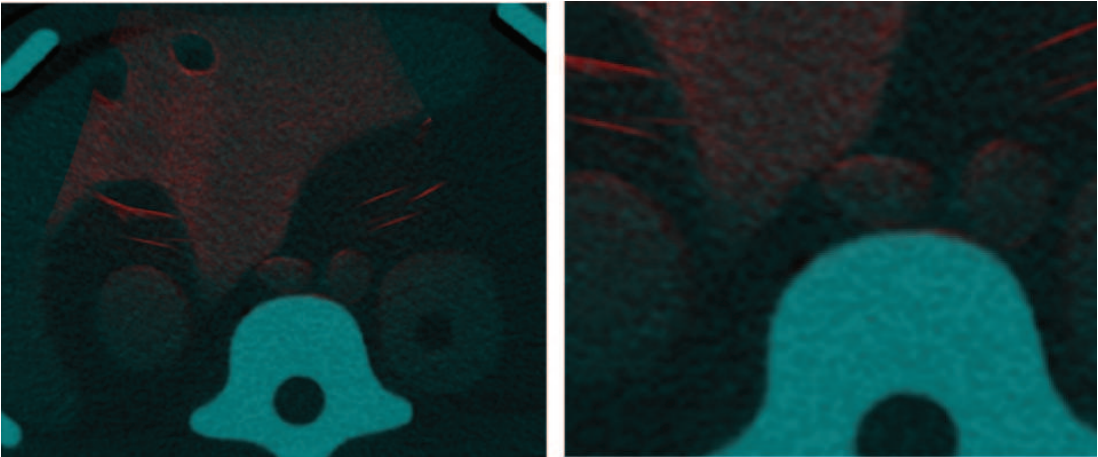


Figure B.27: The two images from g.B.21 overlapped after the T PS registration for $\alpha = 100$

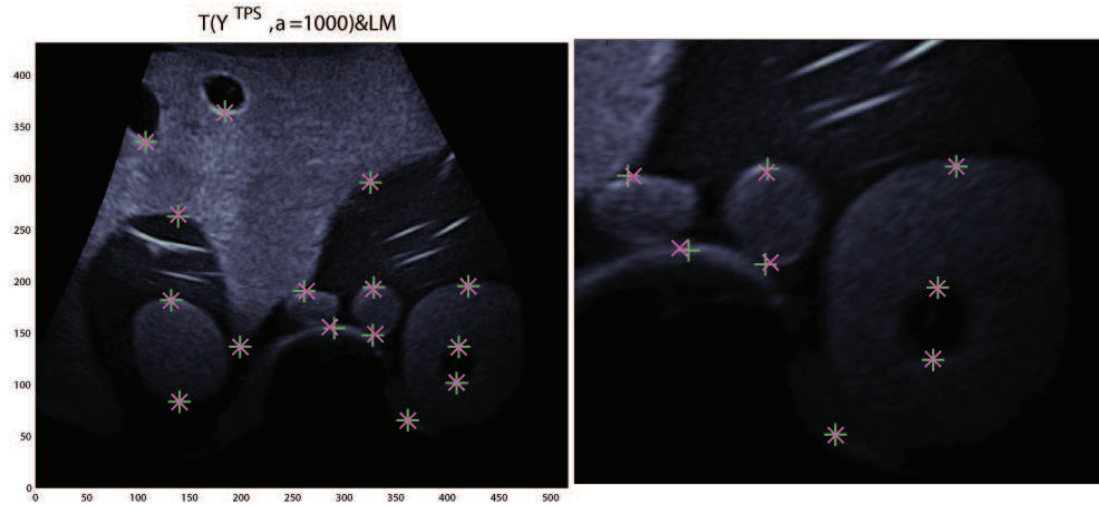


Figure B.28: The target image from g.B.21 registered for $\alpha = 1000$

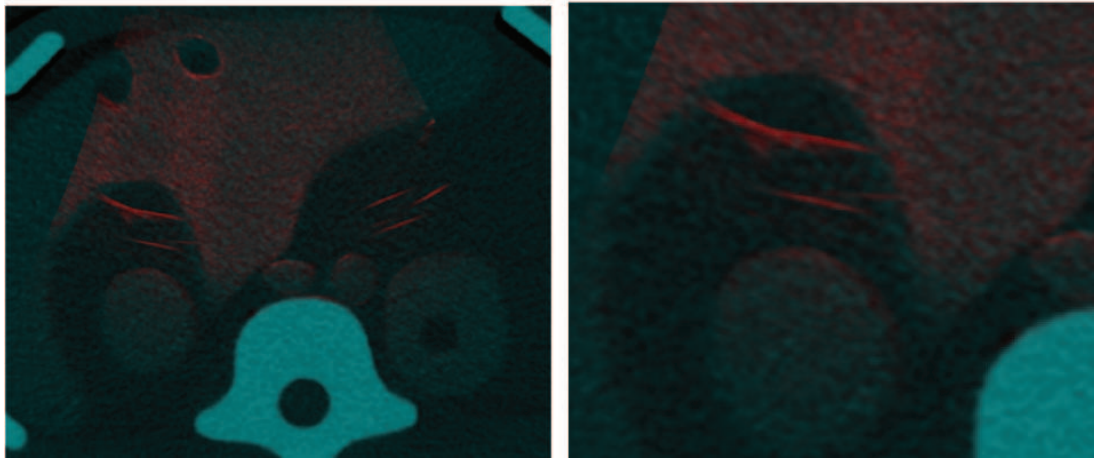


Figure B.29: The two images from g.B.21 overlapped after the T PS registration for $\alpha = 1000$