COMPARISON OF THREE POINT-BASED TECHNIQUES FOR FAST RIGID US-CT INTRAOPERATIVE REGISTRATION FOR LUMBAR FUSION

Benjamin Brat, 260575054, McGill University, Montreal

August 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master in Biological and Biomedical Engineering.

Dedication

À mes parents, France et Éric Brat,

Pour leur amour, leur soutien inaltérable et plus que je ne saurais dire

À mon frère et ma sœur, Sophie et Alexandre,

Pour leur amour et leur sincérité

À ma famille,

Toujours proche, même très loin

À tous mes professeurs qui ont fait tellement plus qu’enseigner

À mes frères et sœurs d’élection,

Qui m’ont soutenu, accompagné, conseillé, émerveillé, et beaucoup plus

Table of Contents

List of Tables

List of Figures

Abstract

Medical image guidance for pedicle screw implantation during spinal fusion reduces the complication rate for patients drastically. Preoperative images require image-to-patient registration to provide accurate information and afterwards, accurate guidance. Automatic registration based on tracked intraoperative ultrasound is a promising method but requires a robust and accurate algorithm. This thesis presents the comparison of three different algorithms that could potentially satisfy all clinical requirements in terms of image-to-patient registration: the Iterative Closest Point algorithm, a more robust variant of ICP and the Coherence Point Drift (CPD) algorithm.

The accuracy, precision, speed, robustness regarding initial misalignment, and robustness regarding noise were evaluated for each algorithm on three datasets of increasing realism. First, they were evaluated on the registration of artificial point clouds in the shape of an ellipsoid. Second, they were tested on the registration of a segmented bone surface of the Computed Tomography (CT) scan of a plastic Sawbones lumbar phantom. Finally, they were assessed on the registration of the segmented bone surfaces of a CT scan of a porcine cadaver with the segmented bone surface of a tracked ultrasound scan of the same cadaver. Accuracy was measured relative to the current gold standard generated with fiducial landmarks. The results demonstrated that even though the three algorithms were equivalent on the two first datasets, only CPD could stay robust and precise on porcine datasets.

The CPD algorithm is a good candidate for automatic, fast and robust intraoperative use. In the future, the testing methods presented here could be expanded to other algorithms.

Abrégé

Le guidage par imagerie médicale pour l'implantation de vis pédiculaires lors de la fusion vertébrale réduit considérablement le taux de complications chez les patients. Les images préopératoires nécessitent un recalage image-patient afin de fournir des informations et un guidage précis. Le recalage automatique basé sur une échographie peropératoire localisée est une méthode prometteuse mais nécessite un algorithme robuste et précis. Cette thèse présente la comparaison de trois algorithmes différents susceptibles de satisfaire toutes les exigences cliniques: l'algorithme Iterative Closest Point (ICP), une variante plus robuste de l'ICP et l'algorithme de Coherent Point Drift (CPD).

L’exactitude, la précision, la rapidité, la robustesse vis-à-vis du désalignement initial et du bruit ont été évaluées pour chaque algorithme sur trois jeux de données au réalisme croissant. Ils ont d'abord été évalués sur le recalage de nuages de points artificiels sous la forme d'un ellipsoïde. Après cela, ils ont été testés sur le recalage avec lui-même d'une surface segmentée de la tomodensitométrie (TDM) d'un fantôme lombaire en plastique Sawbones. Enfin, ils ont été évalués sur le recalage des surfaces osseuses segmentées d'un scanner d'un cadavre de porc avec la surface osseuse segmentée d'une échographie localisée du même cadavre.

La précision a été mesurée par rapport à l'étalon-or actuel généré par des repères fiduciaires. Les résultats ont montré que même si les trois algorithmes étaient équivalents sur les deux premiers ensembles de données, seul le CPD pouvait rester robuste et précis sur les données porcines.

L'algorithme CPD est un bon candidat pour une utilisation peropératoire automatique, rapide et robuste. À l’avenir, les méthodes de test présentées ici pourraient être étendues à d’autres algorithmes.

Acknowledgements

I want to thank my supervisor, Dr. D. Louis Collins, for the opportunity and trust he gave me, for the support and patience he showed constantly and for his openness and accessibility. Further thanks extend to Pr Brandon Xia for his advice and support along the way.

My gratitude to:

- Dr. Vladimir Fonov for having a solution to any technical problem.
- Dr. Simon Drouin for welcoming me and guiding me when I arrived.
- Dr. Houssem Gueziri for his patient help, insights, and joyfulness.
- Pina Sorrini for being helpful, supportive and patient.
- My colleagues in the Neuroimage and Surgical technologies Lab: Anka Kochanowska, Dr Mahsa Dadar, Dr Alonso Cárdenas de la Parra, Dr Azar Zandifar, sMarzieh Tahaei, Dr Phil Novosad, Dr Ryan Sanford, Dr Ian Gerard, Daniel DiGiovanni, Jai Hebel, Victoria Madge and Étienne Léger

Contribution

I am the principal author of this thesis. I am responsible for image analysis, which included part of the image processing stage, quality control, statistical analysis, visualization, and interpretation of the results. Several other persons contributed, specifically:

- Pr D. Louis Collins: Primary supervisor of the project. Provided valuable insight and advice on the data analysis method.
- Dr. Charles X B Yan: former Ph.D. student under the supervision of Pr Collins. Responsible for collecting all the phantom and porcine data.
- Dr. Houssem Gueziri: Post-doctoral Fellow in Pr Collins' laboratory. Provided valuable insight and technical help with the data early preprocessing up to the segmentation stage and aided with the interpretation of the results.

Chapter 1. Introduction

Surgery, and specifically, neurosurgery is a medical practice in constant effort to reduce the risk of lasting damage for the patients. The development of ever more precise tools or more accurate imaging techniques furthers this endeavor. 4 in 5 people will suffer from back pain at some point in their life [1] making it the third most important cause for doctor's visits [2], therefore responsible for millions of dollar in health spending. The importance of reducing as much as possible the long-term unwanted consequences of a spine surgery becomes more than merely the care for patients' well-being, but a public health problem. Spinal fusion surgery, in particular, is performed in cases of spondylolisthesis, scoliosis, spine fracture, dislocation, spine stenosis, spinal tumors, and pseudarthrosis [3]. The spinal fusion surgery is composed in several steps: opening the patient, resecting muscles, insertion of the pedicle screws, fixation of the rods and closing. A decisive moment of several spine surgery techniques is the insertion of metal screws in the vertebral transverse pedicles, used in order to fix rods and plates to the bones. There are two reason for this decisiveness. First, the surgeon has limited information on the orientation and depth of the screw with traditional methods; second the vertebral pedicle is very close to critical anatomical structures such as the spine, blood vessels or nervous branching out of the spine. One in four screws is implanted inaccurately [4], which can have dire consequences on the patients' surgical outcomes. The maximum acceptable error on the screw positioning varies between 1 and 2mm for translation and 5° for orientation depending on the screw size, vertebral level, and anatomical context. Each vertebral fusion necessitates at least 4 screws, which puts every patient virtually at risk of serious long-term damage.

In this situation, Image-Guided Spine Surgery (IGSS) provides the much-needed localization and orientation tools to surgeons in real-time inside the operating room. The use of IGSS devices has been shown to decrease both the number and importance of positioning errors [5]. Those devices collect and present in an efficient way orientation information acquired using multiple sources, whether radiative, magnetic or sound-based. One critical step of some IGSS software is the registration of the preoperative CT scan with the intraoperative anatomy of the patient. Indeed, if multiple imaging techniques are used at different moments, the task of finding the correspondence of position and rotation between the multiple scans might not be a straightforward task. The traditional method to accomplish this is to equip a probe with tracked markers so that the 3D coordinate of the tip of the probe is known, and then, to identify specific landmarks from the scan. Identifying a large number of landmarks manually on each modality at regular intervals is a tedious, slow, and error-prone method, not adapted to the context of the operating room (OR). Besides, the registration accuracy deteriorates with time as the patient may move during surgery, or the reference object used as the origin of coordinates for the patient space may move or be bumped.

In order to avoid these difficulties, several medical device manufacturers propose a large number of intraoperative imaging devices such as O-Arm or C-Arm.[6] The advantages of those methods are on the data treatment, indeed no need to register multiple imaging techniques together if the most precise method is already available in the OR. These devices all function based on X-ray imaging and reconstruction, which implies a more massive radiation dose for the patient and numerous technical difficulties for the operating team. For example, OR staff using an MRI must have all their material be made of a non-magnetic alloy. Similarly, the use of CT scanners on top of forcing to stop the operation to scan the patient requires to remove every metallic components or instrument in the room. Ultrasound (US) scanning has been tested as a worthwhile alternative to X-ray imaging. Ultrasound cumulates many advantages, such as the non-ionizing wavelength used, the relatively low cost of the machine, and the reduced space it takes inside the operating room. Unfortunately, ultrasound also has some drawbacks. In particular, the amount of noise and artifacts on a US signal requires a large amount of preprocessing and processing in order to obtain accurate information automatically from the data. The most uncomplicated information to extract automatically from ultrasound scans is certainly the bone surface because the acoustic impedance difference between bones and soft tissue is three times larger on average than the acoustic impedance between different soft tissues [7]. Such difference makes it easier to localize the interface between the two. In addition, the bone surface is quite distinctive in CT scans, as well. This distinct feature in both imaging methods, gives clear information about the coregistration of these scans as well as the visualization of all the data in the same coordinate space. This second aspect attempts to remove the burden of visualization from the surgeon to the computer.

A conventional, computationally cheap, and fast algorithm able to operate a rigid registration between the bone surfaces on both modalities is the Iterative Closest Point (ICP) algorithm, proposed first by Besl and MacKay in 1992 [8]. It is an expectation-maximization algorithm using a kD tree to identify pairs of closest points and afterwards, a metric which is more ofter than not the Euclidian distance in order to compute an average transformation between the two datasets. Since then, several enhancements to the ICP algorithm were published and implemented. The most popular is probably the Coherent Point Drift (CPD) algorithm, proposed by Myronenko and Song in 2010 [9]. This algorithm is not restricted to medical imaging and has shown excellent results in multiple computer vision applications. Non-rigid registration methods are also developed simultaneously. However, the challenges they present are vastly different, which legitimates them being subject to a separate analysis.

The present work will focus on the comparison of three techniques for the rigid registration of vertebrae between CT and US scans. These include the standard ICP algorithm as it was initially proposed, a more robust version of the standard algorithm where datasets outliers are rejected, and the CPD algorithm.

First, we will present the clinical context, then the main evolution in the approach of this problem in the last few years. We will then describe the tracking structure in which the registration step is embedded, following with the detailed description of the experiments led to compare those algorithms and the results, which we will display, discuss and comment.

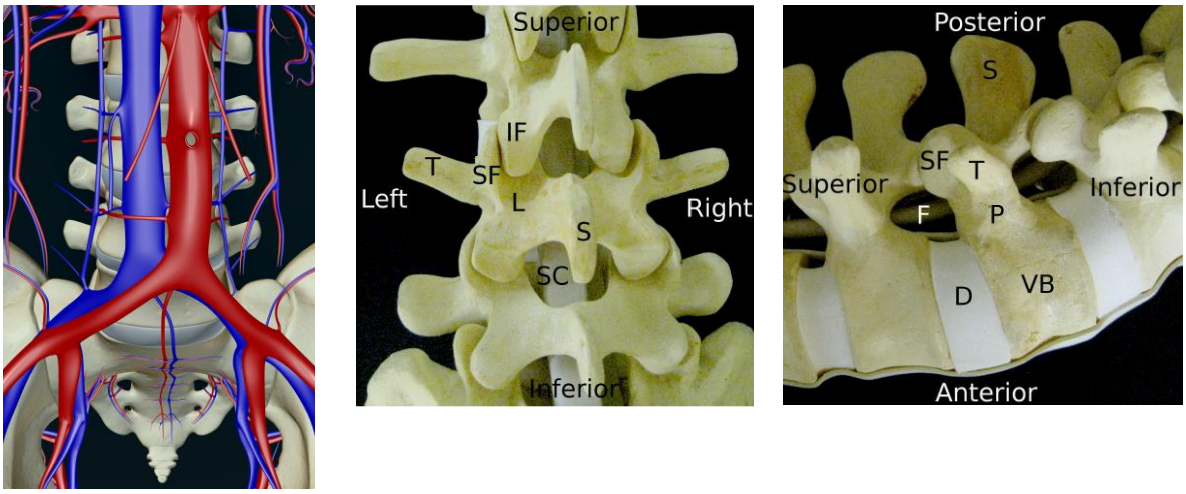
The comparison allowed us to determine that the CPD algorithm is more robust and more precise than the two other algorithms, even in the presence of substantial noise, with a runtime compatible with clinical use. The algorithm is, however, dependent on the segmentation biases, which illustrates the critical importance of this previous step of the image processing. The systematic testing carried in this work could and should be extended to all registration methods in order to legitimate a standard practice that would achieve the best results for each clinical application.

Chapter 2. Background

The improvements in imaging and computing technology have enabled medical practitioners to implement surgery methods that are simultaneously faster, more precise, and more accurate than traditional surgery that is based only on experience, anatomy knowledge, and visual information. One of those transformed fields is image-guided surgery. Classically, surgeons rely on their knowledge of anatomy and physiology, their experience, and the information available in the operating room, such as the visual aspect of the patient or monitored biological quantities, e.g., the heart rate. Some operations were carried without visual information since exposing the relevant structures of the body would create too big a risk for the patient. However, such minimally invasive techniques may have non-negligible detrimental effects on the precision and outcome of the operation. The development of imaging technology has led to an increasing number of techniques available, expanding the amount of information accessible to surgeons and support staff.

2.1. Vertebral anatomy and spinal fusion

One of the operations that benefited from this change is spinal fusion, and specifically, the step of pedicle screw insertion. The vertebrae are bones protecting the spinal cord and supporting the weight of the whole skeleton, as can be seen in Figure 2.1. They are surrounded by muscles, both assuring posture and movement of the body. The spinal cord inside the vertebral canal has nerves branching out between each vertebra. Major blood vessels, the inferior vena cava and the descending aorta run anterior to the vertebral column. The most common surgical access to the vertebrae is from the back or posterior side, allowing the surgeon to approach the spinous processes, the laminae, the facets, and the transverse processes, but not the vertebral bodies or the pedicles. The screws are implanted through the pedicle but enter the bone by the lamina.



a) *b)* *c)*

Figure 2.1 :From Left to Right, a) Anatomical view of lumbar vertebrae with bones in grey and blood vessels in red and blue, image credit: MedicalRF.com/ Alamy Stock Photo b) Dorsal view of a foam phantom of lumbar vertebrae IF: inferior facet T:transverse process SF: superior facet L: lamina S: spinous process SC: spinal canal. c) Sagittal view of the same phantom P: pedicle F: intervertebral foramen VB: vertebral body. [10][11]

During vertebral fusion, several vertebrae are rigidly attached in order to limit their relative movement. This is done by implanting a screw within the pedicle and the vertebral body on both the left and right sides of the successive vertebrae to be instrumented. The screws will then serve to anchor metallic rods. An example of an apparatus mounted on a phantom can be seen in Figure 2.2.

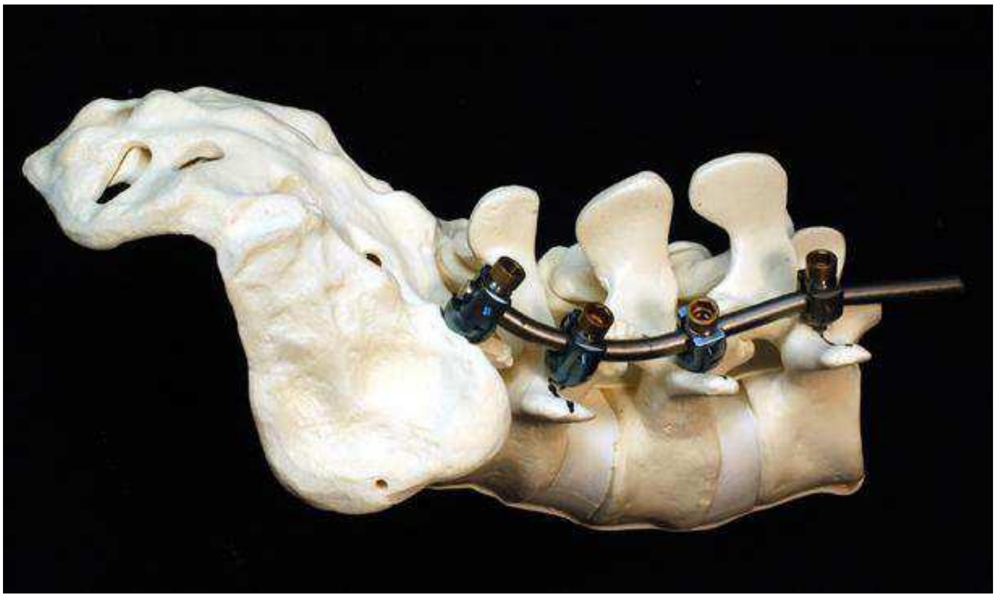


Figure 2.2 : Metallic rods and pedicle screws in situ on a phantom model [12]

Before the advent of image-guided surgery, the common practice was to insert the screws between the superior facets and the transverse processes [10], through the pedicle and through the vertebral body, using resistance feedback as a guide in order to avoid breaching any of the bone walls.

2.2. CT imaging and pedicle screw insertion

Computed Tomography (CT) is an imaging technique used every day in hospitals in order to obtain anatomical information without direct contact with the anatomical structure of interest. At its core, the technique uses multiple X-ray scans of the patient at different angles in the transverse plane in order to reconstruct a 3D volume of the patient where organs that absorb X-rays slightly appear light and organs that absorb X-ray a lot appear dark. One of the main factors for X-ray absorption is tissue density; therefore, bones are the darkest structures. Another X-ray based technique is fluoroscopy, which acquires 30 images per second, allowing for real-time monitoring at the expense of more radiation dose for the patient and blurring factors, both spatial and temporal. [7]

Ultrasound volumes, on the other hand, are acquired using a probe that sends soundwaves with a frequency higher than the audible spectrum, receives the echo back, and computes the time of flight. Knowing the speed of sound in tissues allows for localization of the echo sources. An echo happens when the soundwave hits an interface between two tissues of different acoustic impedance, which corresponds to tissues of different densities. A line of detectors acquires a 2D image. Sweeping the detector perpendicularly to the image plane and aggregating the 2D images constitutes a 3D volume of ultrasound data. Since contrast comes in ultrasound imaging from the tissue density, bones have excellent contrast in US. [7]

Optimal scanning patterns in the specific context of lumbar vertebra scanning was compared by Yan et al. [10]

Preoperative CT is used to choose the screw dimensions and identify the insertion point. Fluoroscopy is used during surgery to monitor the screw penetration and avoid a breach of the vertebral body. Screw misplacement chances during this surgery may vary because many spinal structures are not directly observable, and patient anatomy can sometimes vary significantly. Depending on the direction of deviation of the screw, the misplacement may have vascular, nervous, or muscular consequences on the patient’s recovery. Studies show that the rate of misplacement when the surgery team only refers to preoperative CT scans and fluoroscopy monitoring of the insertion varies between 5 to 50%. [14,15] The high failure rate was attributed to the positioning of the dynamic reference basis (sometimes abbreviated DRO, DRB, DRF), an object that acts as the center of the coordinate system in the patient frame. With the development of medical imaging, the spine surgeons can choose more precisely the insertion point, the angle, and the depth of penetration coherent with the positioning information given by image-guided surgery systems.

2.3. Image-guided surgery

The help that computing power and medical image acquisition could provide to spine surgery has led to the emergence of image-guided surgery systems. Their help is twofold: first they allow visualization of structures and tools that would otherwise be hidden, and second, they allow the combination of visual information from different sources in the same reference frame, for example, the surgical plan, the preoperative CT scan, and fluoroscopy information can be displayed on the same computer screen. Such systems are composed of an optical camera tracking the position of infrared targets, as well as reflective spheres placed on several tools in unique 3D configurations allowing for their identification. [14,16]

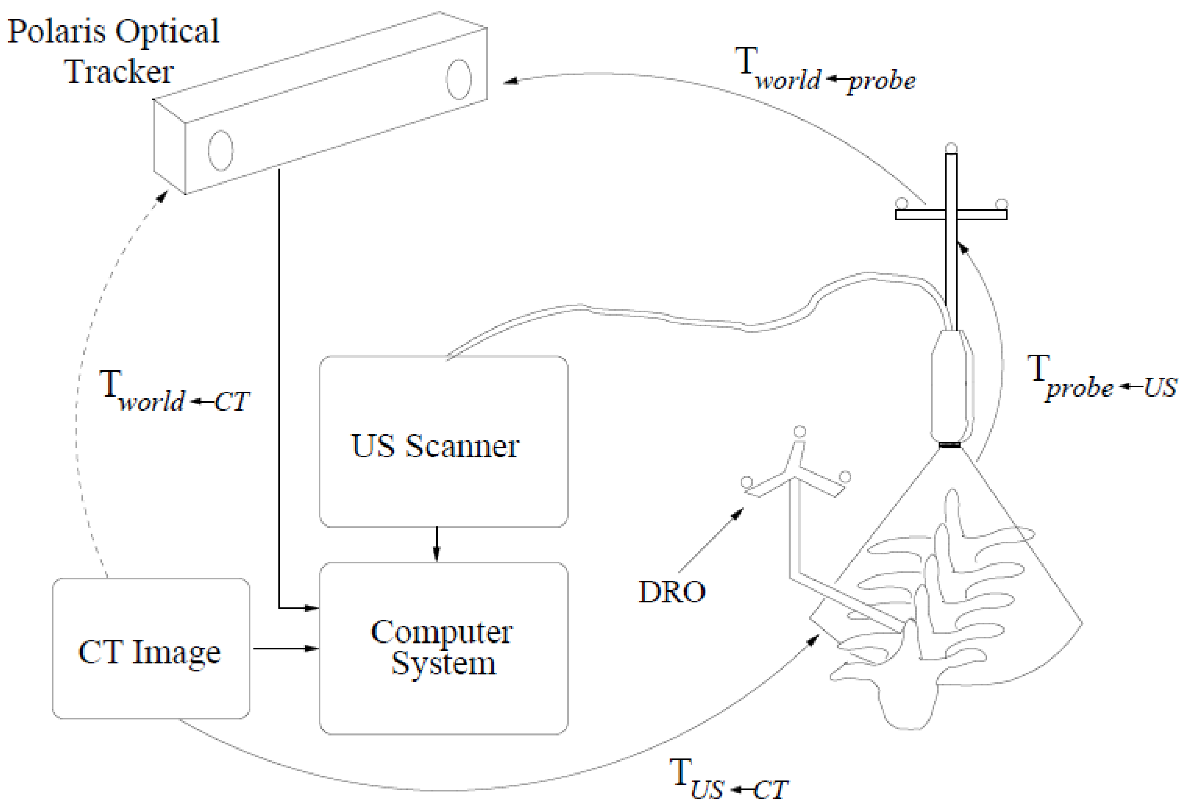


Figure 2.3 : Schematic representation of an ultrasound based IGSS [16]

The camera emits an infrared signal that is reflected by the spheres, allowing for fast and accurate 3D tracking. They have a submillimeter precision laterally and a precision of approximately 1mm deep, even when the object is placed several meters away from the camera, allowing the medical team's movements not to be hindered by the presence of the cameras. One of the targets is chosen as the coordinate reference for the patient space and is known as the Dynamic Reference Object (DRO). Other tools or elements equipped with tracked spheres are the intraoperative imaging devices, surgical tools, and pointers. Commonly, one of these markers is attached to the patient. The calibration of the probe pairs each voxel in the US volume with a coordinate relative to the probe position and orientation. The tracking system can give the position of the scanning probe relative to the DRO. This means that the US voxel coordinates in the patient frame can be computed by the system. The goal of the registration step is to find the relationship between the two frames of reference: the CT scan frame and the US volume frame. This is one of the tasks of the computing system (3DSlicer [17], IGSTK [18], StealthStation [19], IBIS [20], CustusX [21], and others) which integrates all the information, register every image and tracked position together and enables processing and visualization of the result. A schematic showing the elements of a US-based IGSS can be seen in Figure 2.3. The goal of the computer system is to determine the patient-image transform $T_{world \leftarrow CT}$ allowing to associate to each CT scan voxel a coordinate in the physical world, called world frame or patient frame. In order to do so, the transformation $T_{US \leftarrow CT}$ coregistering CT and US voxels shall be determined. One can note that it does not matter in which coordinate frame the US volume is for the registration with the CT data. As the transformations that bring the US volume to the world frame are very well defined by the tracking system and probe calibration, they can also be applied to the CT data.

2.4. IBIS

The present work will focus more specifically on IBIS, the Intraoperative Brain Imaging System [21,23]. This open-source software combines tracked ultrasound and augmented reality in order to avoid deterioration of the patient-image registration accuracy and provides the surgeon with the most visually efficient navigation information. This platform, optimized for spine and brain surgery, provides algorithms optimized for fast GPU computation and intuitive graphic user interface. The platform has been tested both in the laboratory and in clinical use on a significant range of neurosurgical operations: brain shift correction, assistance in tumor resection, neurovascular surgery, pedicle screw implantation, path planning for deep brain stimulation, and epilepsy electrode implantation. [22]

The registration protocol and algorithm for spine surgery are published in Dr. Charles Xiao Bo Yan's doctoral thesis "Ultrasound-CT registration of vertebrae for image-guided spinal fusion surgeries." [12] The present work aspires to extend modestly on the research conducted by Dr. Yan.

The clinical accuracy requirements necessary for image-guided spinal surgery were established by a group of 70 experts. Their conclusions were published by Cleary *et al.* [4]. For segmentation and registration, the image guidance accuracy necessary is 1-2 mm, for a total duration of a few minutes. This means that any point on the patient should be associated with a point on the other imaging modalities scans that is no more than 2 mm away from the exact location in a timespan of a few minutes.

2.5. Patient-image registration

2.5.1. Manual registration

The earliest image-guided surgical systems used manual registration of landmarks on preoperative CT scans. The preoperative scan was segmented beforehand, and the bone surfaces were isolated and organized into a vertebral model. During the surgery, the posterior (paraspinous and subperiosteal) muscles were retracted or dissected to expose the transverse processes.

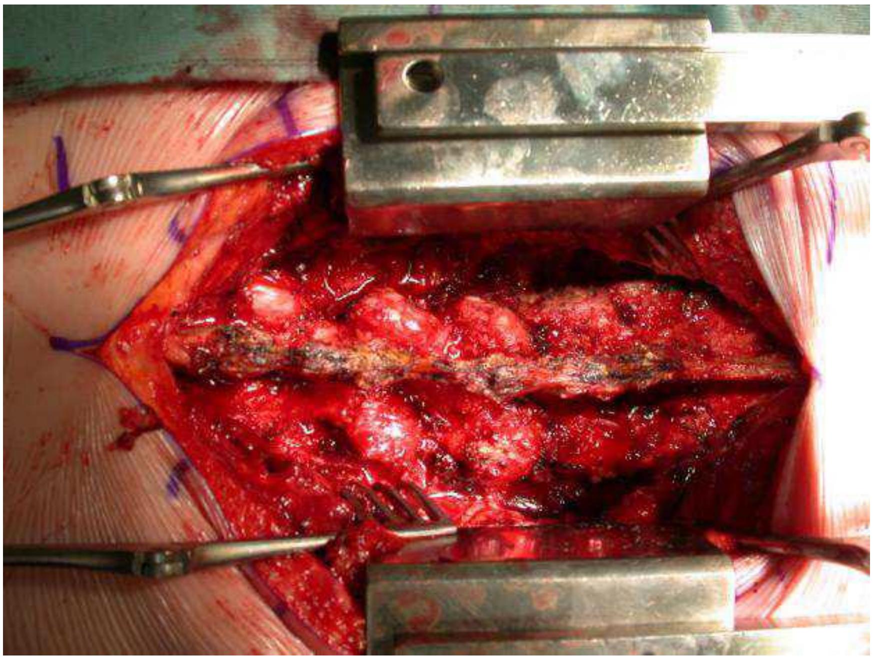


Figure 2.4: Picture from a neurosurgeon point of view after incision and muscles retraction, exposing the vertebral anatomy.

Then, using a tracked pointer, the positions of landmarks on the patient anatomy were paired with their homologous points on the CT scan. Afterward, the system could generate a transformation allowing the registration of both spaces together. Transferring the CT scan in the patient space then allowed simultaneous visualization of tracked tools and anatomical information from the registered images.

Verification of the transformation validity is made regularly during the surgery by checking that pointing to a location on the bone surface with the tracked pointer shows the same location being pointed at in the rendering on the computer screen. This process lasts approximately 5-10 minutes, for average total procedure time, counting registration, verification, and screw insertion, of 20 minutes per vertebra. Using manual registration greatly diminishes the misplacement rate to around 8%[2]. This manual method, however, presents several drawbacks. It significantly increases the operation time, proportionally to the number of vertebrae to fuse, and requires broader access to the bone surface in order to reach the landmarks, requiring the dissection of more soft tissue. This technique constitutes, however, the standard of care in terms of image-guided vertebral fusion, which is why we will use the registration information from this technique as a gold standard in order to compare the three automated registration methods investigated in the present work.

In order to mitigate the drawbacks of the manual method, several intraoperative imaging techniques were developed. The first idea was to acquire the anatomical image directly inside the operating room, removing the intermediary step of correcting a preoperative image to the intraoperative position. The other idea was to use another imaging technique that would more efficiently identify anatomical information, reducing the time required for registration.

2.5.2. Automatic registration

In general, the intra-operative imaging tool is fitted with tracking targets and calibrated beforehand in a way that its images can be registered in the patient space without any manual input. [23] An intraoperative image of the patient is acquired using either 2D or 3D fluoroscopy, Computed Tomography (CT) scanning, Magnetic Resonance Imaging (MRI), or ultrasound (US) scanning. Fluoroscopy or CT share the drawback of exposing the patient (and surgical staff) to increased doses of X-ray radiations. This is not desirable, and even less so for young patients. Intraoperative MRI does not emit ionizing radiation, but it creates other issues. The first one is undoubtedly the excessive cost of the intraoperative MRI scanner. However, it also requires the use of magnetically compatible surgery instrumentation, it limits the positions of the patient during the procedure, as well as access to the surgical field. Compared to these imaging modalities, ultrasound scanning provides numerous advantages. It is economical, maneuverable, fast, and non-radiative. These advantages were interesting enough for Ault and Siegel to publish a proof of concept as early as 1995. [24] It has, however, the main drawbacks of poor image quality, making it sometimes difficult to interpret the US images. Some techniques were developed recently by Hacihaliloglu et al. in order to detect tissue surface accurately in ultrasound images. [25–27] It has, nonetheless, been regularly used in conjunction with preoperative CT or MRI imaging for navigation during surgeries of the brain, the pelvis, the prostate, the liver, or the kidneys. The present work focuses on ultrasound-based image-guided surgery of the spine. These systems function on the same principle, as explained previously in section 2.3. From the surgical team's point of view, in a quick gesture with the ultrasound probe already used in medical practice and a few seconds of computation, the preoperative CT images would be aligned to fit the patient's real-time position and grant the benefits of image-guided surgery without the time loss or the supplementary dissection required for manual landmark-based registration.

There are two categories of ultrasound-CT registration methods: point-based registration and intensity-based registration. Point-based registration requires preprocessing and segmentation, which means extracting the bone surface points in both the CT scan and the ultrasound. [28] With the point clouds obtained, several techniques can be used to find the transformation to align the two point-clouds by minimizing the distance between them, e.g., Iterative Closest Point algorithm. Intensity-based registration, on the other hand, does not require precise surface segmentation. A method circumventing the segmentation would, for example, involve computing an image that would be similar in intensity level and distribution to a reconstructed ultrasound scan based on the pre-operative CT data and then finding the optimal transformation between the pseudo ultrasound scan and the scan acquired during the operation. Intensity-based registration techniques usually have considerable runtimes but are, in general, more robust. [29] The point-based registration techniques, on the other hand, require segmenting the images, and the optimization process is susceptible to being stuck in a local minimum of the optimization function, therefore computing a suboptimal transformation. The present work does not cover segmentation, but several articles publishing techniques of automatic segmentation of vertebrae in CT images report runtimes compatible with a surgical context. [30] Due to the considerable interest of the domain in the research community, many articles were published over the years. As mentioned by Oliveira and Tavares [31] in their own 2014 review, several reviews of the domain were compiled.

Chronologically, the first review about image registration was published by LG Brown in 1992. [32] In 1998, Maintz and Viergever wrote one of the first review papers on medical image registration. [33] Their group published another review some 20 years later. [34] They confirmed the shift they predicted from extrinsic registration to intrinsic registration methods, as well as insisting on two problems of the field still unsolved, validation of registration methods, and translation in clinical practice. Hill et al. [35], Oliveira and Tavares [31], and Mani and Arivazhagan [36] wrote other medical images registration reviews, also insisting on the major change towards intensity-based methods and the facilitation that constituted computing power improvements for the field. Hill et al. dedicated a large amount of their review to intensity-based metrics, and specifically, those who originate in information theory, while Oliveira and Tavares insisted more on having a presentation as broad as possible, addressed to researchers unfamiliar with the field. Mani and Arivazhagan have written a very detailed conclusion describing a large number of possible developments and future research directions. Tjardes and Shafizadeh [13] summarized the state of the art in Image-Guided Spine Surgery, identifying the directions that future techniques shall improve, and explicitly insisting on the interaction between clinicians and software as a source of error. Schumann [37] summarized the state of the art in ultrasound-based registration for Computer-Assisted Orthopedic Surgery, advocating for fully automated processes and more validation studies *in vivo,* while Langlotz and Nolte [38] published a technically oriented review on CAOS insisting on the importance of long follow-up time studies. Alam and Rahman [29] detailed the contributions in intrinsic registration techniques for medical images, concluding that landmark-based registration techniques are the most appropriate to medical images, due to their reliability and stability at the cost of processing time. Ferrante and Paragios [39] surveyed slice-to-volume medical image registration, which are techniques circumventing the 3D reconstruction step. Finally, Noble, Navab, and Becher [40] reviewed the advances in ultrasonic image analysis concerning surgery.

The first feasibility study of computer-assisted pedicle screw fixation was published by Amiot et al. [41] in 1995 and achieved a 4.5 mm accuracy, with a success rate of around 85%. Ault and Siegel [24] published the concept proof of ultrasound-based registration with the same accuracy. Helm et al. [15] focused on the results of spinal navigation studies done between 2000 and 2014, showing that on the 12622 pedicle screws placed using varied state-of-the-art methods, the success rate reached 96.8%, and investigating the reason why these techniques are not more widely spread in daily practice. Brounstein et al. [42] carried automatic non-rigid point-based US-CT pelvic registration using Gaussian Mixture Model, achieving a mean surface accuracy of 2 mm in seconds.

Meyer et al. [43] applied the ultrasound-guided surgery system presented in an earlier section with augmented reality to anesthesia. By combining a preoperative patient model and tracked needles monitored using ultrasound, they achieve lumbar facet injection therapy with an error inferior to 1mm, compared to 5 mm on average with classic methods, demonstrating improved clinical results when using a US-guidance system.

Finally, Yan et al. [10] carried automatic intensity-based US-CT lumbar vertebrae registration and achieved an error inferior to 1mm on the landmarks localization of a foam lumbar phantom and error inferior to 2mm on porcine lumbar cadavers.

2.6. Literature review of point-cloud registration methods

The research in point cloud registration algorithms stemmed from the seminal article by Besl and McKay [8] in 1992. As mentioned by Pomerleau [44] in his review on robotic applications of point-based registration, between 1992 and 1993, ICP was applied in four domains: object reconstruction, non-contact inspection, autonomous vehicle navigation and medical, and surgery support. The three former domains are not in the scope of the present work, but several results from these domains will be discussed for their interesting algorithmic development. As the first feasibility study for spine application was only published in 1998 [41], many algorithmic developments that were initiated beforehand were not specific to US-CT rigid registration for lumbar pedicle screw implantation. In their article, Besl and McKay propose a simple expectation-maximization algorithm, which will be detailed in section 3.1.

Just like medical images registration review papers, several review papers gathered and organized the available knowledge on point-based registration from an algorithmic approach. The first article that built on the Besl and McKay ICP algorithm is Chen and Medioni [45], who introduced point-to-surface matching. Champleboux et al. [46] introduced the idea of outlier rejection in their implementation of the Levenberg-Marquardt (LM) optimization with curve-fitting algorithm and octree decomposition in order to compute point-to-surface distance efficiently.

Fitzgibbon [47] introduced a new method for registration based on Besl and McKay algorithm but optimized by adding an LM fitting implementation. This LM-ICP is faster, more robust, and multi-purposed than the initial algorithm. This differs from Champleboux et al. [46], cited above, by the implementation of outlier error mitigation by using a more robust kernel in the error function.

Zhang [48] introduced the idea of decomposition into free-form curves and implemented outlier rejection efficiently in the ICP algorithm.

Rusinkiewicz and Levoy [49] published a review specifically on the variants of the Iterative Closest Point algorithm, comparing their convergence speed. The most efficient algorithm they presented could register two groups of 3D points in less than 0.1s, which is more than enough for the clinical application discussed here.

Ma and Ellis [50] published results showing that a more robust metric than the Root Mean Squared Error, such as Tukey's bi-weight error, could increase the algorithm robustness. This version of the ICP algorithm is called the ICP-Tukey. This algorithm seems to handle outliers better than the original ICP but performs worse on cleaner datasets. It was tested on digital phantoms of tibia, femur, and vertebra.

Pomerleau et al. published a review of point cloud registrations over the last twenty years [44] as well as, the same year, a comparison of several ICP variants on real-world 3D volumes considered challenging due to their complexity and variety called the ETH set[51].

Dr. Magnusson's doctoral thesis [52] presented the 3D-Normal Distributions Transform (3D-NDT) method for registration. Although the method is primarily intended for mobile robotics, it was successfully applied to medical images. The fundamental advantage of this method is to transform the point clouds into smooth surfaces allowing for classical optimization methods, Newton's method, for example, to converge. Indeed, most numerical optimization methods depend on the smoothness and curvature of the cost function. The comparison between 3D-NDT and standard ICP showed that in the context of robotics, 3D-NDT is faster and less dependent on the initial estimate; however, ICP showed a more robust convergence.

Petricek and Svoboda [53] used the same ETH 3D volumes dataset as Pomerleau [51] to compare the results of several methods: generalized ICP, 3D Normal-Distribution Transformation (NDT), Fast Point-Feature Histograms, 4-point congruent sets and their keypoint-based method. For overlap bigger than 75%, which is the clinical situation, their method achieved 0.03 rad rotation error and 0.58m translation error (compared to at least 1m for every other method) in 14s on average. This is much longer than ICP but still fits the clinical requirements, given that the distances considered there are far more significant by one or two orders of magnitude compared to a medical context.

Chui and Rangarajan [54] formulated the registration problem in a Bayesian probabilistic framework, allowing for optimization using a probabilistic algorithm. Specifically, they developed the Gaussian Mixture Model (GMM) algorithm, which is less sensitive than ICP to outliers and noise.

Xie *et al.* [57] compared the efficiency of ICP and the Expectation Conditional Maximization algorithm developed by Horaud et al. [58] on a minimal number of points from a pelvic scan. They showed that even with sets of 50 points, the ECM algorithm could correctly register them together 90% of the time without noise and 37.5% of the time with 1mm Gaussian noise.

Yang *et al.* [62] tackled the problem of local minima by developing a Branch and Bound approach in order to start the ICP algorithm close to a local minimum, allowing to compare them and selecting the one having the lowest error. The authors report convergence of the algorithm in seconds with submillimeter root mean square error, but the code was not tested in a medical image context.

Myronenko and Song [9] published a variant to the GMM algorithm called Coherence Point Drift (CPD). It involves forcing the point-clouds centroids to move coherently. For rigid registration, it allows them to determine a closed-form solution to the maximization equation. Combining the Bayesian framework with topological considerations allows this method to achieve robust registration even in the presence of noise, outliers, or missing points in seconds. The algorithm was validated on synthetic datasets as well as cardiac ultrasound scans. Wang *et al.* [63] added a correction to the CPD algorithm. The original method had a manual parameter *w* which measured the proportion of outliers and noise in the data. They provided an automatic method to determine the best value for this parameter.

Maier-Hein *et al.* [64] presented works done in the goal of considering inhomogeneous and anisotropic noise in the ICP algorithm, A-ICP. Indeed, the original version of the ICP algorithm postulates homogeneous Gaussian noise, which is not always the case. For example, due to the difference in accuracy between lateral and depth of infrared Time-of-Flight cameras, the noise structure of tracking information is certainly anisotropic. Their method involves integrating the use of covariance matrices as point weighting. Those covariant matrices are computed using either Principal Component Analysis or the Voronoi area of points on a surface.

Billings *et al.* [66] proposed their probabilistic version of the ICP algorithm called the Iterative Most Likely Point algorithm. This method allies the advantages of a probabilistic framework with its noise robustness and the efficiency of not having to carry every pairwise point matching. Moreover, the generalized noise model allows taking into account anisotropic noise. Their algorithm is compared to ICP, Generalized ICP, CPD, and A-ICP on both hip-bone and femur, for multiple conditions of noise and outliers.

Cao *et al.* [67] published a variant of the standard ICP algorithm called the Closest Point Transform, which relies on the use of a closest-point map on a grid instead of matching points in real-time. This significantly reduces the computation time of the algorithm at the cost of some interpolation. They tested their method on head scans in CT and MRI, achieving results comparable to volume-based methods, relying on mutual information.

Jian and Vemuri [69] reformulated the registration task in a probabilistic framework by introducing a Gaussian Mixture Model. This led them to establish a framework where the relationships between multiple GMM implementations are clarified. They compared their GMM based algorithm to CPD and LM-ICP, among others, on 2D and 3D datasets. However, no test was carried out on medical images. No substantial difference in the convergence of the three algorithms was established.

Evans *et al.* [71] investigated the use of a tracked stereovision camera system in order to co-register CT scans and stereo camera scans of porcine vertebrae. They showed that a TRE of around 2.5mm was possible to achieve; however, this method would require much more investigation.

Moghari and Abolmaesumi [72] published a variant of the Unscented Kalman Filter transform to achieve registration. Originally, Unscented Kalman Filters compute probabilistic estimates of hidden variables when provided with some non-linear observables, as well as an initial position. This framework allows for the inclusion of many different noise or bias sources. The kD-tree algorithm would pair the points from both datasets together, and the Kalman Filter would compute the most probable transformation, taking into account the different sources of anisotropic noise. This algorithm was tested on pelvic CT scans and scaphoid CT bone phantom.

Brendel *et al.* [28] published in 2002 a method for the registration of 3D CT and ultrasound scans for the spine that uses intensity for surface-to-volume registration. They mention achieving a Target Registration Error of 2mm in translation and inferior to 1 degree in rotation. This gradient-based method of forward ray-tracing for CT is similar to the one currently implemented in IBIS.

King *et al.* [73] address the shortcomings of using ultrasound data in the operating room, which are a low signal-to-noise ratio and several artifacts. In order to avoid a large amount of preprocessing, that US scans must go through, they use the preoperative scans in order to establish probability density maps and simulate ultrasound data. The contour from preoperative images transformed into ultrasound images using probability density maps is then registered with the real ultrasound images. This technique was tested on heart scans and used for respiratory motion correction.

Over the years, numerous ameliorations were proposed to the original ICP algorithm, but no thorough comparison of every candidate was carried for any specific application. The literature regularly features LM-ICP, CPD, or 3D-NDT as a comparison basis besides ICP, but the relative efficiency of those methods is not established as it depends strongly on the constraints of the registration task. The present work initiates this investigation in the specific case of lumbar vertebrae US-CT registration.

Chapter 3. Methods

In the present study, we will compare the robustness and efficiency of three common point-based registrations algorithms for lumbar vertebrae rigid registration. The three methods compared are firstly the initial ICP algorithm by Besl and McKay [8], secondly, a more robust version of the same algorithm which includes an outlier rejection step [48] and, finally, the Coherent Point Drift algorithm. [9]

3.1. ICP algorithm

ICP is a method to co-register two or more point-based representations of an object in the same coordinate space. It requires decomposing the surface of the objects into point-clouds, then finding the linear transformation that minimizes the Euclidian distances between each pair of points. This algorithm is an example of an Estimation-Maximization method in the sense that it proceeds in two steps at each iteration, first matching points from the different representations together based on closeness, then computing the optimal transformation that minimizes the cost function, which is the sum of squared distances between the matched points. This algorithm is fast, simple, and computationally cheap, as the lower bound on the number of iterations is O(n*log n) where n is the number of points. [8] However, it presents a few limitations: it cannot efficiently handle large transformations; it requires a reasonable starting estimate, and it is also sensitive to outliers and noise.

The ICP algorithm implementation requires the following steps for registration of points X_{N*3} corresponding to the vertebral surface as seen in CT with the second set of points Y_{M*3} corresponding to the vertebral surface as seen in ultrasound [8]:

1. Initialization with the fixed point-cloud of ultrasound segmented bone surface X_{N*3}, the moving point-cloud of CT segmented bone surface $Y_{M*3}$and an initial estimate of the rotation matrix R_0 and the translation vector t_0 that would register the two point clouds together, usually $R_0 = I_{3*3}$ and $\overrightarrow{t_0} = \vec{0}$. Two parameters should be specified: $threshold$ which represents the threshold value of the cost function that is considered acceptable and $MaxIterations$ which represents the maximal number of

iterations the algorithm should go through before stopping if no satisfactory error level was attained. The error function use din ICP is the root mean squared error of the two set of points. In order to study the influence of the number of iterations on the convergence, the threshold on the error value was fixed at 0.00001mm, an unrealistic value that would make the condition on the threshold always false.

2. Apply the transformation to the CT scan points, where "i" serves as index of the number of iterations, "j" serves as index of the elements of Y and Y', "n" serves as index of the elements of X and "^{T}" is the transpose operation:

$$Y'^T = R_i * Y^T + [\vec{t_i} \quad \vec{t_i} \quad \ldots]_{3*M} \tag{1}$$

3. Find the closest CT point from the set of points X for each ultrasound point y'_j using a kd-tree algorithm [76]:

$$(\forall y'_j \in Y')(\exists x_n \in X) \ \ s.t \ (x_n, y_j') = kdtreematching(X, Y') \tag{2}$$

4. Find the transformation that would minimize the root mean square error between the two datasets:

$$(R_{i+1}, \overrightarrow{t_{i+1}}) = \underset{x_n \in X}{argmin} \left\{ err_i = \sqrt{\frac{1}{N} * \sum_n \left\| x_n^T - R_{i+1} * y_j'^T - \overrightarrow{t_{i+1}} \right\|^2} \right\} \tag{3}$$

5. Then repeat from step 2 until either the maximal number of iterations is attained, or the root mean square error between the two datasets is inferior to a fixed threshold: if $(err_i \geq threshold) \cap (\, i < MaxIterations)$ then repeat from step 2 with $(R_i, \vec{t_i}) \leftarrow (R_{i+1}, \overrightarrow{t_{i+1}})$

Having an idea of the shape that the RMSE curve should take as a function of the transformation parameters would help interpret the experimental results. Because we are considering only rigid registration, the distances between points inside the point cloud do not vary. This implies that the position of the barycenter, or center of mass, of the cloud relatively to each point is constant. Studying the rigid movement of a point cloud around another one can, therefore, be simplified as the movement of their barycenter, combined with the rotation of the point clouds around their barycenters. Let us simplify the problem for an instant and consider that the point clouds do not rotate around their barycenter. Then, the only movements become translations of the point clouds along an axis and rotation around an axis that does not pass through the barycenter. For each matching pair of points of the two datasets, a translation increases the distance between them linearly, and therefore the Root Mean Squared Error (RMSE). Supposing that the error is minimal for no transformation, as is the case with our datasets, the RMSE is then simply the amplitude of the translation, in absolute value. In the case of rotation around an external axis, there is only one plane defined by the line perpendicular to the axis passing through the first barycenter and the line perpendicular to the rotation axis passing through the second barycenter. On that plane, the movement of one barycenter relative to the other is a rotation. The distance between two points on a circle is given by equation (4). Those represent substantial simplifications of the general case but give us an element of comparison for the RMSE curves. Therefore, the shape of the Root Mean Squared Error (RMSE) for this extreme case of perfect fit would be the shape every other less fitting datasets' RMSE should tend to. In that case, the error described in equation (3) boils down to the distance between the two points. Analytical solutions are easy to derive as the distance between two points on a line if they were translated away from each other, or on a circle, if they were rotated away from each other:

$$RMSE_{ideal}(t,\theta) = \begin{cases} |t| & for\ a\ translation\ t \\ R\sqrt{2*(1-\cos\theta)} & for\ rotation\ of\ angle\ \theta\ and\ radius\ R \end{cases} \quad (4)$$

An excellent approximation of the two curves, presented in Figure 3.1, is the V-shape of the curve corresponding to translations. Now, considering the volume in which the points are clustered around the barycenter, as the matching between points of the two datasets is based on the smallest distance, then when the two datasets are close to registered together, their spatial recovering already decreases the value of the RMSE. Even if one dataset is not optimally placed, the fact that the other dataset overlaps with it will contribute to the decrease of the RMSE value close to the minimum. This explains why the RMSE curves of experiments are more likely to be shaped like a U than like a V.

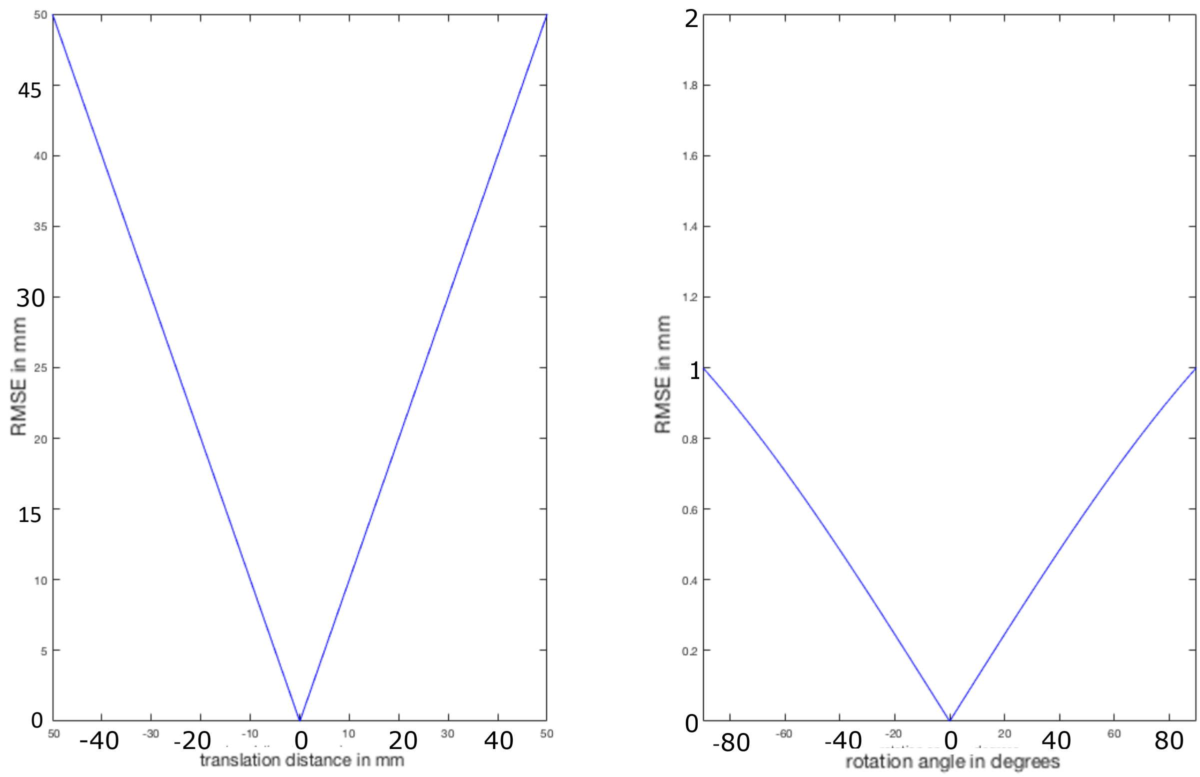


Figure 3.1: Ideal shape of the RMSE for a translation on the left, and for a rotation for a circle of radius R=1 on the right.

The non-linear trend of the curve for the rotation case is negligible for the range of angles coherent with a clinical setting.

We observe the following limitations while applying the ICP algorithm. Outliers are points that do not correspond to the surface of the object. They can be situated arbitrarily far from the actual surface. Noise is a perturbation that assigns a slightly wrong position to points of the surface due to errors in segmentation or partial volume effects, for example.

It must be noted that the noise and outliers considered in the registration step are residual from the initial imaging noise. Indeed both US and CT have built-in or reconstruction based error correction features. For CT scanning, iterative reconstruction algorithms allow for less noisy reconstruction, and collimators placed before the detector weed out the outliers. For US, compound scanning can be used to reduce speckle in a scan, and harmonic imaging uses harmonics of the soundwave sent in order to increase the resolution and reduce the eventual artifacts. [7] After those imaging steps, the only possible error correction steps are algorithmic, and operated by the computing system.

The classic ICP algorithm does not factor in either noise or outlier correction. The cost function used, however, contains an averaging operation, which makes the contribution of

Gaussian white isotropic noise tend towards 0 for a large number of points. From a theoretical standpoint, this amounts to considering that the noise distribution is purely Gaussian and isotropic so that taking the mean in the computation of the error will reduce the error amplitude due to the noise. As mentioned before, this is probably not true in practice, due to the anisotropic noise of the tracking devices to name one cause, which is why ICP is vulnerable to this kind of error. These artifacts may manifest themselves in a non-monotonous character of the optimization function. Several local minima may be present. The algorithm might mistakenly converge to one of them instead of the global minimum corresponding to the best solution. This leads to a suboptimal transformation, which minimizes the cost function locally but may not minimize the Target Registration Error (TRE) at all. The TRE is the root mean squared distance between the target points in each point cloud. It allows for the computing of the error specifically around zones of interest instead of globally. [77] It is the error metric that will be used here. Usually increasing the number of points in the point-clouds helps smooth out some local minima. A substantial part of the following chapters will be dedicated to this problem of data quality. These problems lead to corrections of the ICP algorithm in an attempt to mitigate those error sources. Many articles mentioned in chapter 2 devised an original answer to the problem of outliers and noise. Each of these solutions has merits and drawbacks, which depend strongly on the application context. The noise sources, the acquisition method, or the nature of the data investigated all influence the noise structure and the outliers' bias. Therefore, each practical application requires testing the multiple solutions available in order to determine the most adequate. One of the methods used for this that keep the main elements of the original algorithm is the rejection of the worst match.

3.2. ICP algorithm with outlier rejection

An elegant way to take care of outliers in the ICP algorithm is the rejection of the worst matched pairs. In substance, if a proportion of outliers p is present in the point-clouds, then those outliers will have the most significant error value when the algorithm is close to the solution. Their effect on the algorithm is not the first cause of error. However, as it is impossible to guess (estimate) *a priori* how close the algorithm is to the solution, removing the proportion p of the matched pairs of points that are the furthest apart from each other will diminish the algorithm sensitivity to this source of error, and thus increase its robustness. In practice, this addition to the classic ICP algorithm translates into inserting a step between 3 and 4 in the above description that eliminates a proportion p of the pairs of points having the largest distance, based on the assumption that these points are outliers:

$$\begin{array}{ccc} if\ dist(x_k, y_j') \in p^{th} percentile & \Rightarrow & exclude\ the\ pair\ (x_k, y_j') \\ of\ the\ distribution & & of\ the\ optimization\ step \end{array} \quad (5)$$

The drawback of this approach is that removing points from the optimization step will increase the effect of noise on the cost function. The fundamental difference between outliers and noise is that a noised point is a point from the zone of interest that was moved around by random effects, whereas outliers are not at all indicative of any zone of interest.[44] The expected value of noised points converges to the real position, unlike outliers which bias the data. Indeed, even in the ideal case of isotropic Gaussian noise, the averaged noise amplitude only tends to the expected value of 0 in the limit of a large number of points. This creates a trade-off on the value of p. Either p is too small, and outliers' biases are the dominant cause of errors, or p is too big, and the noise becomes the dominant cause of errors. Tests were made for different values of p in order to determine this threshold for every dataset tested in the experiments below.
The MATLAB implementation used for ICP registration and robust ICP was published by Kjer and Wilm as part of their bachelor thesis. [78]

3.3. Rigid CPD algorithm

The Coherent Point Drift algorithm tries to tackle the limitations of the ICP algorithm in a radically different manner by using a probabilistic framework. The rigid CPD algorithm considers one of the point-clouds as a list of centroids of a Gaussian Mixture Model (GMM) probability distribution, and the other point-cloud is the set of data points generated by this distribution. In this context, the algorithm maximization step consists of finding the transformation that maximizes the likelihood of correspondence between the GMM and the data points. For iUS-CT registration of vertebrae, the GMM centroids are used to represent the points of the bone surface from the CT scan, and the data points are the intraoperative ultrasound (iUS) points. This algorithm also ensures the topological coherence of the point-clouds by reparametrizing the centroid location. By including a uniform distribution term corresponding to noise and outliers, this algorithm shows remarkable robustness regarding data degradation, while keeping only one free parameter w, representing the proportion of improper points in the data. The Fast Gaussian Transform algorithm [79] computes more efficiently a sum of exponentials, decreasing the number of operations from $O(MN)$ to $O(M+N)$ for the computation of the matrix of matching probabilities in the following CPD algorithm, where M is the number of CT points and N the number of iUS points. According to the creators of the CPD algorithms, the proportion of improper data points (noise, outliers, missing data) can be bigger than half the dataset, even 0.7 for certain situations, and still allow for a robust registration of the two datasets, while ICP outlier rejection rate cannot be bigger than half the dataset maximum, due to fit underdetermination. Finally, being much more robust than LM-ICP on several datasets[9], the CPD algorithm is less sensitive to initial estimates of the transformation.

This algorithm requires the singular value decomposition of an M by N matrix, however, which is computationally intensive. Another downside of the CPD algorithm is that its cost function is not the RMSE measured in mm but the Negative LogLikelihood (NLL) of the distribution, which has no physical unit and therefore does not allow for a direct estimation of the localization error. Finally, the algorithm depends on the initial parameter w, which is difficult to estimate a priori.

The CPD algorithm is implemented as follows [9]:

1. Initialization with the fixed point-cloud of ultrasound segmented bone surface X_{3xN}, the moving point-cloud of CT segmented bone surface $Y_{3xM}$and an initial estimate of the rotation matrix R_0 and the translation vector t_0 that would register the two point clouds together, usually $R_0 = I_{3x3}$ and $\overrightarrow{t_0} = \vec{0}$. The scaling parameter s is 1 for every dataset studied here. The proportion of outliers and noise in the data is w, it is analogous to the parameter *p* used for outlier rejection in the robust ICP algorithm but the original typology is used in order both to refer to the original article more easily and not to confuse it with the probability matrix P or its elements p_{mn}.

2. Compute

$$\sigma^2 = \frac{1}{3MN}\sum_n \sum_m \|x_n - y_m\|^2 \tag{6}$$

3. Estimation step: Compute $P = [p_{mn}]$ with

$$p_{mn} = \frac{\exp(-\frac{1}{2\sigma^2}\left\|x_n - s * R_i * y_m - \vec{t_i}\right\|^2)}{\sum_m \exp(-\frac{1}{2\sigma^2}\left\|x_n - s * R_i * y_m - \vec{t_i}\right\|^2) + (2\pi\sigma^2)^{1.5}\frac{w}{1-w}\frac{M}{N}} \tag{7}$$

4. Maximization step: compute

$$N_p = sum(P) \tag{8}$$

Compute the center of mass of X: μ_X, the center of mass of Y: μ_Y and subtract these center of mass from the two datasets, in order to center them on the origin

$$\mu_X = \frac{1}{N_p}X^T P^T \vec{1}_{M*1} \tag{9}$$

$$\mu_Y = \frac{1}{N_p}Y^T P \vec{1}_{N*1} \tag{10}$$

$$\hat{X} = X - \overrightarrow{1_{N*1}} * \mu_X^T \tag{11}$$

$$\hat{Y} = Y - \overrightarrow{1_{M*1}} * \mu_Y^T \tag{12}$$

5. Compute

$$A = \hat{X}^T P^T \hat{Y} \tag{13}$$

And find the singular value decomposition

$$A = USV^T \tag{14}$$

Then

$$C = \begin{bmatrix} 1 & 0 & 0 & \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \\ & \cdots & & |UV^T| \end{bmatrix} \tag{15}$$

$$R_i = UCV^T \tag{16}$$

$$s = \frac{tr(A^T R_i)}{tr\left(\hat{Y}^T * diag\left(P * \overrightarrow{1_{N*1}}\right) * \hat{Y}\right)} \tag{17}$$

$$\vec{t_i} = \mu_X - sR_i\mu_Y \tag{18}$$

$$\sigma^2 = \frac{1}{3N_p}\left(tr\left(\hat{X}^T * diag\left(P * \overrightarrow{1_{N*1}}\right) * \hat{X}\right) - s * tr(A^T R_i)\right) \tag{19}$$

6. The new moving point cloud is

$$Y' = T(Y) = sYR_i^T + \overrightarrow{1_{M*1}} * \vec{t_i}^T \tag{20}$$

The correspondence probability of the pair (x_n, y_m) is p_{mn}.

The negative log likelihood of this iteration is

$$NLL_i = -\sum_n \log\left(\sum_m p_{mn} + \frac{3}{2}N * log(\sigma^2)\right) \tag{21}$$

7. Repeat from step 3 until either the maximal number of iterations is attained, or the negative Log Likelihood between two iterations is inferior to a fixed threshold:

if $|(NLL_i - NLL_{i-1})/NLL_i| \geq threshold) \cap (i < MaxIterations)$

(22)

Then repeat from step 3 with $(R_i, \vec{t_i}) \leftarrow \left(R_{i+1}, \overrightarrow{t_{i+1}}\right)$

The MATLAB implementation used in this work is the official implementation called "pcregistercpd"[80].

3.4. Datasets

Three datasets with increasing complexity and realism were used to test the algorithms. These include ellipsoids, images of a vertebral phantom, and finally images of a porcine phantom.

3.4.1. Ellipsoids

In order to verify the code implementations of the three algorithms presented above, ellipsoids were simulated using three different axis lengths (a, b, c) and the random generation of 2 angles, φ and θ. φ was varied between -π and π , representing the azimuth, and θ between -$\frac{\pi}{2}$ and $\frac{\pi}{2}$, representing the altitude. These variables can be used to define 3D data points on the surface of an ellipsoid, using the following parametrization:

$$\begin{cases} x = a * \cos\theta * \cos\varphi \\ y = b * \cos\theta * \sin\varphi \\ z = c * \sin\theta \end{cases} \tag{23}$$

Different sets of random numbers (θ, φ) create different samplings of the same ellipsoid. One of the samplings, A, is then transformed using a rotation, a translation, and some additive gaussian isotropic noise of a specified standard deviation. The other, B, is left as is. The three algorithms' task is then to register together A and B. In the subsequent tests, the value used for (a, b, c) is $(100\,,50\,,30)$ because without being too large the difference between each value is large enough to prevent the infinite number of equivalent solutions due to axial symmetry of oblate and prolate ellipsoids. Even then, two opposed global minima exist. The convergence to any of these minima would be satisfactory as this test is only to confirm the validity of the convergence algorithm. Similarly, unlike for the next experiments, no target point is chosen, and no TRE is computed, as it would not have any anatomical significance, this evaluation is more focused on the convergence stability of the algorithms for simple 3D shapes.

3.4.2. Sawbones phantom

3.4.2.1. Sawbones acquisition

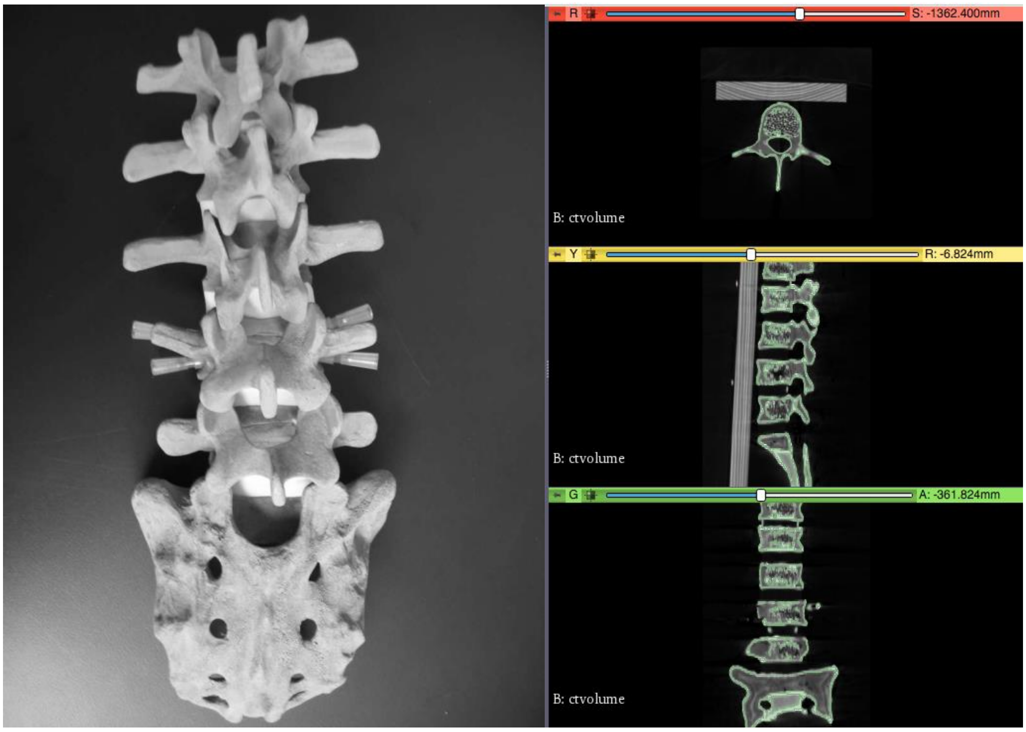


Figure 3.2: Sawbones radiopaque lumbar phantom, picture on the left, sections of the CT scan on the right

The Sawbones Radiopaque Lumbar Phantom 1352-39 (Sawbones USA, A Pacific research Company, Vashon islands, WA) that can be seen in Figure 3.2 is a phantom of the human lumbar vertebrae made of foam. The surface of the phantom is painted with a radiopaque coating to simulate cortical bone. The phantom was attached to a wooden plank to avoid movement and imaged using a CT scanner (Picker International PQ6000) at the Montreal Neurological Institute and Hospital according to the spine neurosurgery protocol. The voxel size of the 3D scan is 0.35 x 0.35 x 2 mm^3 with slices perpendicular to the longitudinal axis. This means that the diagonal of this voxel has a length of 2.02mm. This means that an error of localization of 1 voxel on this CT scan can lead, at worst, to an error of 2mm. The Sawbones phantom images (as well as the cadaver images described below) were acquired by Dr. Charles X B Yan during his doctoral work. More details are found in his doctoral thesis. [10]

3.4.2.2. Sawbones phantom preprocessing

The vertebral bone surface of the Sawbones CT volume was then segmented manually using 3DSlicer [17] in order to obtain as clear a result as possible. The primary tool used was thresholding based on the scan intensity. The removal of several artifacts, such as the spheres used for infrared tracking, was completed by manual erasing.

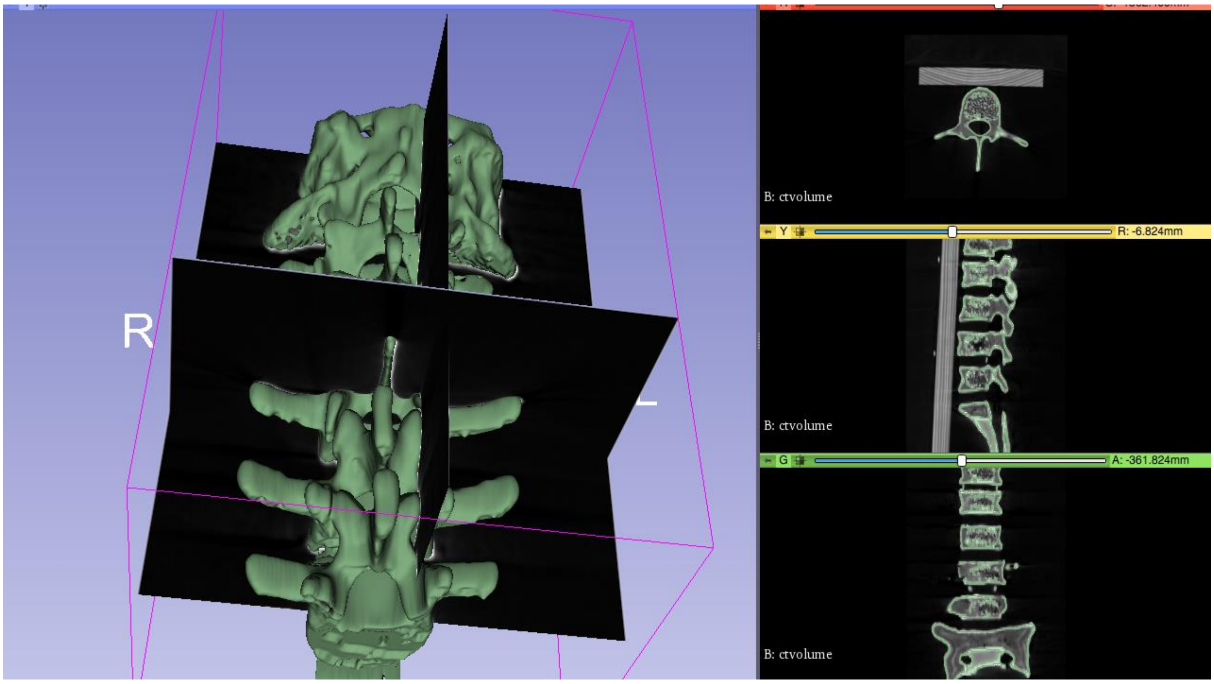


Figure 3.3: Left, segmented CT volume of the Sawbones phantom. Right, sections according to the 3 anatomical plans, from top to bottom, transverse section, sagittal section, frontal section.

The segmented volume visible on the left of Figure 3.3 is from a file containing three lists: an indexed list of vertices for which each row is of the pattern "v *space* [x-coordinate] *space* [y-coordinate] *space* [z-coordinate]" , a list of faces which is just the indices of the three vertices of each triangular face for which each line is of the pattern "f *space* [first index] *space* [second index] *space* [third index]", and a list of the normal vectors to each face for which each line is of the pattern "vn *space* [x-coordinate] *space* [y-coordinate] *space* [z-coordinate]". A point-cloud model of the phantom can be constructed by taking the first list of the file and removing the letter "v" at the beginning of each line. The first vertebra, L1, was manually cut out of the full 507397 spine points transversally at the level of the intervertebral disc, by selecting every point above the transversal plane passing through the middle of the disc. This does not account for the lowest part of the inferior facet of the vertebra but is not critical in order to test the convergence of the algorithms on CT scans. Two points were manually selected at the estimated position of entry of a pedicle screw to be the targets with which the TRE will be computed. The targets are the points of the volume for which positioning is of clinical importance. The localization error on these points will lead to an error on the pedicle screw implantation. Therefore instead of computing the RMSE for every point of the volume, the medical practitioner is only interested in the error the software tolerates on the target points. For the Sawbones phantom, those points can be seen in purple on Figure 3.4.

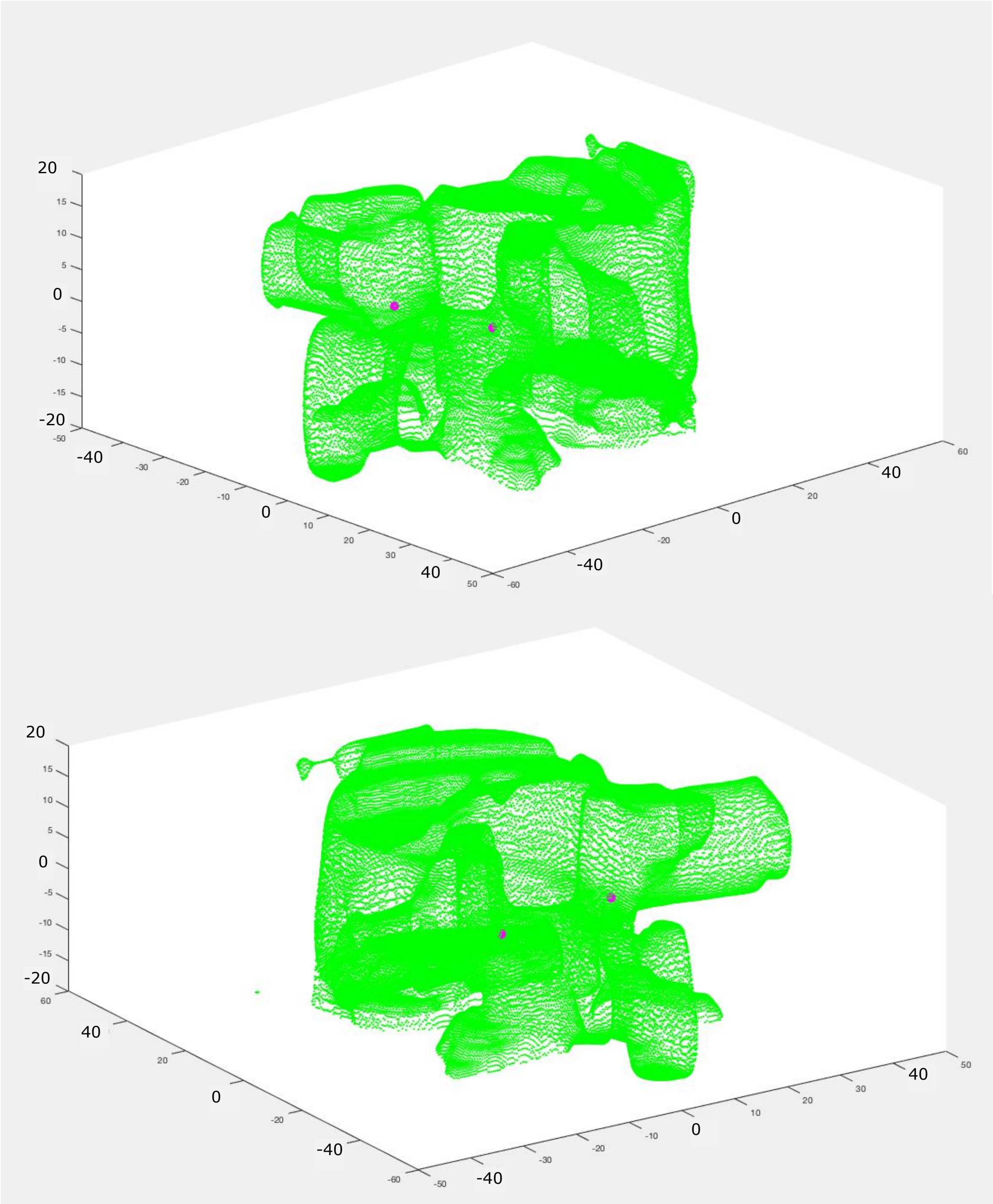


Figure 3.4: Full point cloud of the Sawbones phantom L1 scan with Targets in purple

The CT scan of the Sawbones phantom was therefore transformed into two lists: a list of 71508 points on the L1 bone surface, which large size is due to the fine mesh necessary for a precise segmentation, as well as a list of 2 target points. In order to avoid repetition, two small subsets X and Y of the list of points are randomly selected for each registration. An example of such subsets is presented in Figure 3.5.

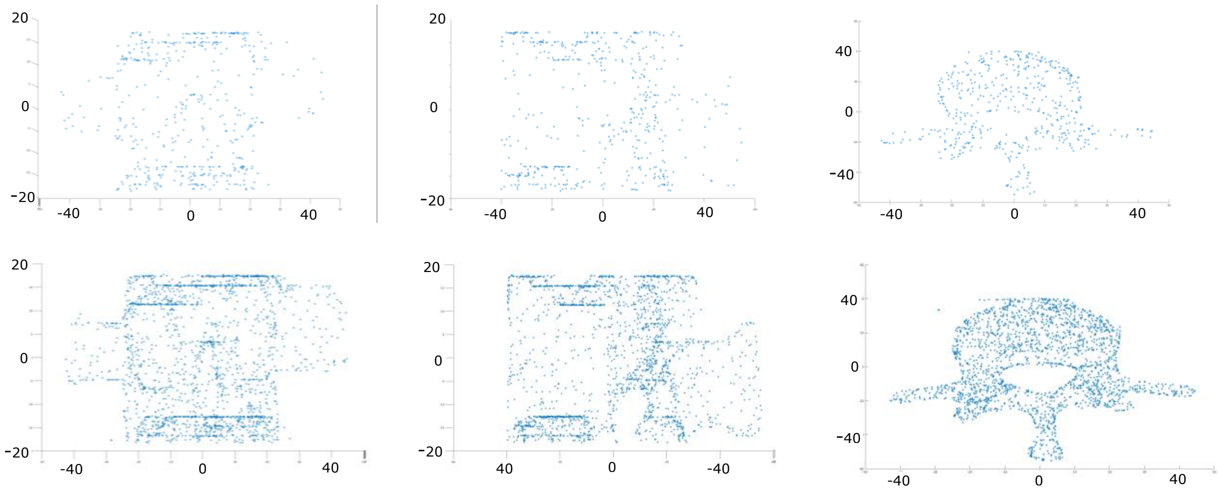


Figure 3.5: Top row, subsampling of 500 points of the L1 Sawbones phantom CT scan, from left to right coronal, sagittal and transverse view. Bottom row, subsampling of 2000 points of the same scan, same order of views.

The number of points necessary in both datasets, as well as the number of iterations necessary for convergence and the rejection rate for robust ICP giving the best results, are investigated beforehand as described in 3.5.2 below. First, a rotated, translated and noised version of both the point-cloud and: the target points are created

$$\begin{cases} Y' = R_{Y\to Y'} * Y + t_{Y\to Y'} + noise(0, \sigma^2) \\ Target' = R_{Y\to Y'} * Target + t_{Y\to Y'} + noise(0, \sigma^2) \end{cases} \quad (24)$$

The rotation and translation parameters are the same for both, and the noise is isotropic Gaussian of standard deviation 2mm. This corresponds to the error of the localization of 1 voxel. The goal of the the algorithms is then to register the original point-cloud X onto the modified point set Y'. This test allows the measure of a Target Registration Error by comparing $Target'$ with $R_{X\to Y'} * Target + t_{X\to Y'}$ using the classic Euclidean distance:

$$TRE = \sqrt{\sum_{Targets} \left\| Target' - R_{X\to Y'} * Target - t_{X\to Y'} \right\|^2} \quad (25)$$

The Sawbones phantom is a good approximation of the bone structure of the human vertebrae from an imaging perspective. The coating and surface texture make it a very faithful representation of a real vertebral bone surface. Indeed, the use of plastic phantoms is widespread in the literature for multiple anatomical situations implying bone elements, e.g., hand [81], pelvis [72], head [22]. However, for a real patient, the spine is surrounded by tissues of different compositions (tendons, ligaments, and muscles), shape and physical properties, which means that the phantom does not replicate the actual patient environment. This would be a significant difference from the clinical case if ultrasound scans were used, but only the CT scan of this phantom was used for this experiment. Ultrasound scans of the plastic phantom submerged in water provide a background cleaner than the clinical reality. The ultrasound data are, therefore, not representative of what can be found in the OR. For this reason, the Sawbones phantom main interest here was to test the registration algorithms on CT data in a "perfect scenario," i.e., for a shape more complex than ideal regular ellipsoids but with point-sets that come from the same surface. Even though the two point-clouds X and Y described above come from the same initial CT scan, the random subsampling makes the point matching step non-trivial, and the noise makes the optimization step more difficult, as would the registration of two independently generated datasets.

3.4.3. The gold standard and porcine cadaver acquisition

The third dataset consists of images of porcine cadaver vertebrae L1 to L6 scanned both using CT and ultrasound, with fiducials implanted into order to be able to compute a gold standard fiducial based manual registration. Fiducials are markers integrated in an image as a reference or measuring scale. On the ventral side of the cadaver, before imaging it in CT, four pipettes' tips were implanted in such a way not to disturb the dorsal side of the scan. Those pipettes' tips were the base on which the fiducials were mounted. For CT, four other pipettes' tips with 4mm diameter steel balls stuck inside were secured in the base. Metal has a high contrast in CT because it absorbs x-rays. [7] For US, the four steel ball bearings in their tips were replaced by reference fiducials whose outward-facing side is centered where the ball bearing was. Having these point's positions in the CT scan and the US scan, the transformation that maps one to the other $T_{US \leftarrow CT}$ is easy to determine.

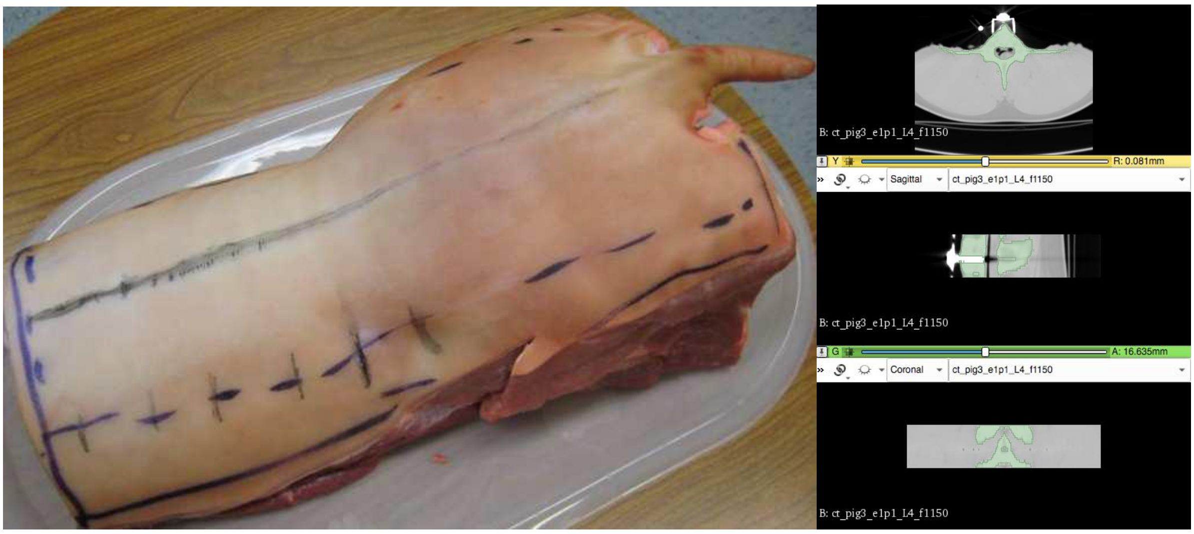


Figure 3.6: Picture of the porcine cadavers prior to dissecting and sections of the porcine's CT scan. A steel ball bearing acting as fiducial can be seen on the transverse and sagittal section

The lumbosacral region of a 60-kg porcine cadaver was acquired at a butcher shop. The lumbar section of the porcine spine is quite similar to the human lumbar spine, both anatomically and functionally. [82] Thoracic vertebrae of the pig, on the other hand, have much longer spinous processes as they have to support the effort of much larger muscles in the pig. This makes porcine lumbar vertebrae a reputed and frequently used biological model for spine instrumentation techniques. The CT scanning of the porcine cadaver was carried in the same way as described above for the plastic phantom with the slight difference that landmarks were implanted.

The gold standard for rigid registration of ultrasound and CT volumes for the porcine cadaver was obtained by implanting markers before both imaging. Markers, which pairing allows us to compute the transformation between the two images. The Oxford English dictionary defines the gold standard as "A thing of superior quality which serves as a point of reference against which other things of its type may be compared "[83], while the definition of ground truth is "A fundamental truth. Also: the real or underlying facts; information that has been checked or facts that have been collected at source."[84] The gold standard does not correspond to the absolute reality, which is traditionally called the ground truth. The ground truth would be the exact movement of each vertebra relative to one another as well as the soft tissues' non-linear behavior. This is beyond what could realistically be accessed here. The gold standard, on the other hand, represents the best available clinical method under reasonable conditions [85]. The four markers used are 4 mm diameter steel balls mounted inside plastic posts made of pipette tips. In order not to interfere with the image acquisition, they were implanted anteriorly. They appeared as bright spheres on CT scans. After that, the markers were replaced by reference fiducials, which position was easy to compute using the Polaris tracking system and a pointer. Because both markers and fiducials were attached to the same implanted base, and because each fiducial centroid corresponds to the marker's center, the position of the fiducials can be quite precisely estimated in both CT and US coordinate system, and the registration parameters are obtained. This manual landmark-based transformation is called the gold standard, and that is the transformation where the automatic point-based algorithms should converge.

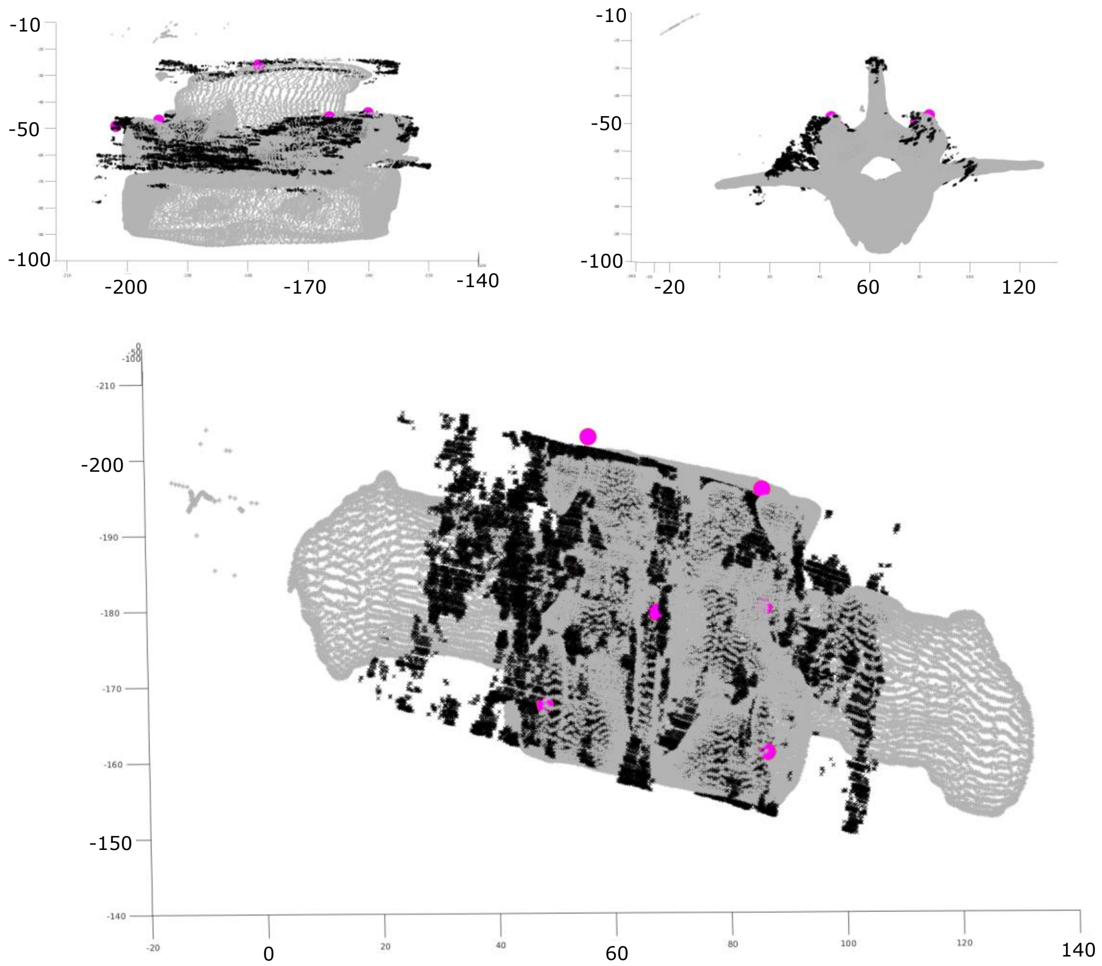


Figure 3.7: In reading order, sagittal, transversal and coronal view of the L3 porcine vertebra point clouds, CT points in light grey, US points in black and targets in purple.

Using these landmarks enables the computation of a gold standard to compare our algorithms.

After these preparations, the CT and ultrasound datasets were acquired and preprocessed.

3.4.4. Porcine cadavers preprocessing

The ultrasound scans were acquired in conditions corresponding to open dorsal surgery. A dorsal midline incision was cut, and soft tissues were dissected or retracted, until most of the posterior surface of the vertebral process was visible, in order to recreate the surgical cavity. The cavity was filled with saline solution (0.9% NaCl in distilled water), allowing for an ultrasound imaging medium. Before data acquisition, the tracked ultrasound probe was calibrated using a Z-bar phantom as described by Comeau *et al*. [86]

The porcine cadaver was scanned using a Philips-ATL HDI 5000 ultrasound system with a multi-frequency phased array probe (4-7 MHz), while maintained entirely underwater. The voxel size of the ultrasound volume is between 1 x 1 x 0.5 mm^3 and 2 x 2 x 0.5 mm^3, depending on the depth. This makes the diagonal of each voxel between 1.5 and 2.87 mm long. A localization error of 1 voxel will, therefore, create an error of at worst 1.5mm close to the probe and 2.87mm far from it. The tracking of the ultrasound probe was computed using reflective spheres attached to the probe and a Polaris infrared camera (from Northern Digital Inc., On., Canada). The probe was swiped from the superior side of the vertebra to its inferior side, with the probe tip posterior to the spinous process, as shown in Figure 3.8.

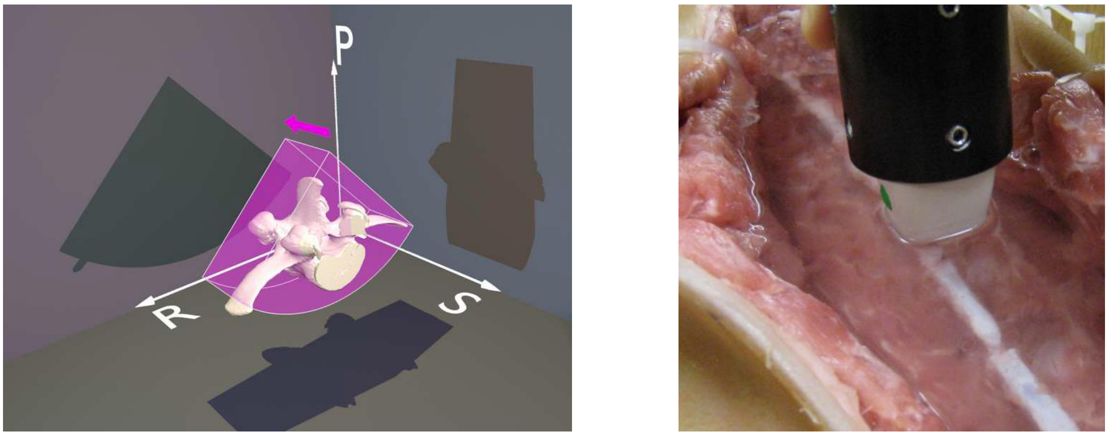


Figure 3.8: ultrasound scanning pattern and picture of the water-filled cavity during acquisition of porcine data

Unlike the Sawbones CT scan, the ultrasound scan was not made as a whole, but each vertebra was scanned separately. Again, the US scans were segmented using 3DSlicer, and the first tool used was intensity thresholding, followed by manual cleaning in order to isolate a bone surface as clear and as thin as possible with reasonable certainty. Surfaces that were not part of the vertebrae were removed from the segmentations, but noise due to residual soft tissues was still present at the end of the manual segmentation process. Another possible approach could have been to use the CT scans in order to segment a US bone surface coherent with the CT bone surface segmentation, as the segmentation process is not under investigation here. This point will be developed extensively in Chapter 5.

Next, several transformations were made in order to coregister US datasets and CT datasets, and have their center of gravity sit at the origin of the coordinate basis. The CT point-cloud and the US point-cloud were registered together into the patient space for each vertebra using the Gold standard transformation $T_{world \underset{gold}{\longleftarrow} CT}$. Indeed, both IBIS and 3DSlicer apply the transform $T_{world \leftarrow US}$ automatically to the US volume and segmented US volume based on the tracking information transmitted by the Polaris tracking system and the probe calibration. Therefore the gold transform, instead of mapping CT voxels with US voxels as one could expect while looking at Figure 2.3, maps CT voxels to world space. Afterward, the point-clouds centroids were translated at the origin. The same translation that centered the ultrasound dataset was applied to each US target point. The target points are chosen in the patient space after CT and US datasets were registered using the gold transform. This implies that any transformation applied to the targets would increase the TRE. For every simulation, three lists of points were necessary: a random subsampling of the US points, a random subsampling of the CT points, and finally, the target points.

Each simulation consists of transforming the CT and target points, and have the registration algorithms correct for this modification using those three lists.

The final number of points in each dataset used throughout this thesis can be found in the Table 3.1.

Table 3.1: Number of points in the segmented datasets used in the present work

dataset	Number of points in the full US scan	Number of points in the full CT scan
Sawbones L1	-	71508
Porcine L1	55288	78936
Porcine L2	60100	81040
Porcine L3	58486	82910
Porcine L4	50823	82512
Porcine L5	48655	87924
Porcine L6	55966	70940

3.5. Experiments

3.5.1. Experiment 1: Algorithm verification

The ellipsoids were used only to verify the three algorithms' convergence. Therefore, the primary indicator of success was the monotony of the cost functions. If the algorithm converges, it is necessary that the quantity minimized by the algorithms, the RMSE for ICP, and negative loglikelihood for CPD, should decrease with each iteration. This is not always the case as some algorithms use multi-starts or other techniques in order to overcome local minima of the cost function. This is not the case here; therefore, we should expect cost function evaluations to have decreasing value across the whole registration process.

3.5.2. Experiment 2: Sawbones phantom

The Sawbones phantom is more anatomically realistic than the ellipsoids, and robustness tests were carried out on this dataset.

Before starting with robustness testing, several parameters need to be chosen: the number of points in both CT and US datasets, the maximal allowed number of iterations that should achieve convergence, and, finally, the value of the outlier ratio used by robust ICP. In order to do so, datasets of 250, 500, 1000, 2000, and 5000 points each were perturbated using only one of the six parameters of rigid registration, in particular rotation around the x-axis, and 2mm Gaussian noise was added to each data point. Then, the three registration algorithms are run, five times per value of the rotation angle, in order to average differences not related to the registration step. The resulting TRE were checked for 40 and 80 iterations. The combination of the number of points in the subsampling and number of iterations with the lowest average TRE value, *N*,* and *i**, is the one that will be used for all the remaining tests. For those simulations, an arbitrary rejection rate of 1% was chosen. Finally, the outliers rate was tested by simulating only robust ICP for around 0% outliers, then 1%, 2%, 5%, 10%, and finally 15%, with 2mm Gaussian noise, *N** points in each dataset, over *i** iterations. For each of these percentages, a scatter plot of 500 random registration tests is computed, showing the final TRE as a function of the initial TRE. The optimal value of the outlier rejection rate, *r**, is the one that minimizes the final average TRE of the 500 tests. These three parameters, *N**, *i*,* and *r*,* are then used for the robustness testing.

In the third set of experiments, we want to evaluate how sensitive to each transformation parameter the three algorithms are. Only one transformation parameter was varied at a time over a broad range of values, (-20°,20°) for rotation angles and (-20 mm, 20 mm) for translation distances, by steps of 2°/2mm. The range of 20° and 20mm amplitude is broader than what would be considered clinically relevant. This experiment aims to determine each method's sensitivity to the initial position - for example, using increasing values of the rotation around the anteroposterior axis before registration. Then, for each transformation, the three algorithms register the two datasets together. Computing the RMSE and the TRE after the registration allows us to estimate the threshold for which each method stops converging.

The fourth test evaluates the sensitivity of the three algorithms to initial conditions. A subsample of *N** US points and *N** CT points is chosen once. Then 500 random registrations are applied successively to the target points and CT points. The TRE is computed before and after registration is done. For each test, there is, therefore, the initial TRE, the TRE after registration by ICP, the TRE after registration by robust ICP, and the TRE after registration by CPD. This creates three scatter plots of 500 points each, where the abscissa of each point is the initial TRE, and the ordinates are the values of final TRE, with different colors for the different algorithms. The most robust method is the one for which the scatter plot ordinates has the smallest standard deviation. Although it may not be the most precise. As the same subsampling is used for every test, local minima of the cost function could maybe trap the algorithms every time. Changing the subsampling of the datasets would hide the presence of plateaus due to local minima of the cost function.

3.5.3. Experiment 3: Porcine cadaver data

The tests carried on the porcine cadaver dataset were similar to those carried on the Sawbones data. First, the optimal parameters for the number of points *N**, number of iterations *i*,* and rejection rate *r** are determined. The numbers of points tested are 250, 500, 1000, 2000, and 5000 points for 40 and 80 iterations, with a rejection rate *a priori* of 1%. Then using *N** points and *i** iterations, 500 different random registrations are carried with rates of 0.001%, 1%, 2%, 5%, 10%, 20%, 30% and 40%.

Robustness tests were carried as described earlier using the optimal values computed above.

First, robustness regarding the transformation parameters was evaluated. Each one of the six parameters of a rigid transform were varied in the interval (-20°,20°) or (-20mm, 20mm) from the gold standard transformation. For these values, registration was carried, and the initial RMSE and the three final TRE (one per registration method) are analyzed in regard to the transformation amplitude. The RMSE allows visualization of the cost function of the ICP and robust ICP algorithm in units of mm at the beginning of the registration. The final TRE comparison shows the parameter amplitude maximum for which each algorithm still correctly registers the two volumes together. A robust algorithm would have a TRE curve relatively flat for every parameter, with possibly some random variations. If the TRE increases drastically, it would mean that the algorithm could not correctly register a transformation of this amplitude. This process was repeated for L1 to L6.

The last test consists of evaluating the robustness regarding the initial transformation. A subsampling of *N** CT points and *N** US points are chosen at random. Five hundred random transformations are generated in an interval of ±20° and ±20mm and applied to the targets and CT points. The US and CT points were then registered together by the three algorithms. For each test, there are four results: the initial TRE, the TRE after registration by ICP, the TRE after registration by robust ICP, and the TRE after registration by CPD. This creates three scatter plots of 500 points each, where the abscissa of each point is the initial TRE, and the ordinates are the values of final TRE, with different colors for the different algorithms. The parameters used were *i** and *r**. The width of the distribution is a direct measure of the robustness of each algorithm. This process was repeated on the US and CT scans of vertebrae L1 to L6.

The same subsampling was kept for each of the 500 tests in order to integrate the presence of local minima to the analysis. From a technological standpoint, comparing the three algorithms by themselves makes more sense in order to determine their differences, all things being equal. However, from a clinical standpoint, no practitioner is interested in the general success rate of their method over hundreds of independent attempts, their only requirement is that the patient that is currently in need of medical attention is correctly treated by the method. In that sense, keeping the same subsampling for every test supposes that the same exact same patient would be treated multiple times to the algorithms. Analyzing this stochastic distribution will indicate the chances for each patient to be well treated. Analyzing the distribution of this test with a randomized subsampling would give us an indication of the proportion of patients well treated. Keeping the same subsampling integrates the potential shortcomings of earlier parts of the process (that led to this specific set of N^* points) and allows us to evaluate the accuracy distribution for one patient individually. An example of this test made with 500 different subsampling will be presented in Chapter 5.

Finally, the runtime was compared for each method, in order to make sure that the algorithms are coherent with the clinical criterion and do not exceed a few minutes of computation.

Chapter 4. Results

4.1. Results of Experiment 1

The qualitative results of ellipsoid registration are presented in Figure 4.1 and Figure 4.2. Figure 4.1 shows the initial relative positions of both datasets, while Figure 4.2 shows the datasets after each algorithm coregistered them. The blue graph corresponds to classic ICP, the green graph corresponds to robust ICP, and the red graph corresponds to CPD.

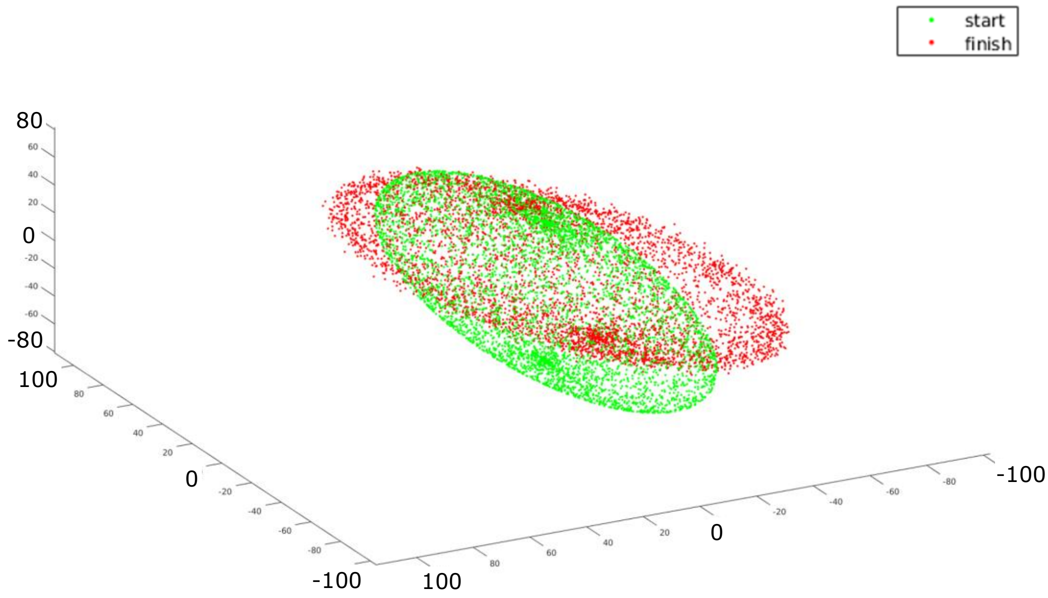


Figure 4.1: example of an initial positioning of ellipsoids before registration. In green, initial ellipsoid. In red, transformed noised ellipsoid. The green set is then registered to fit the red one.

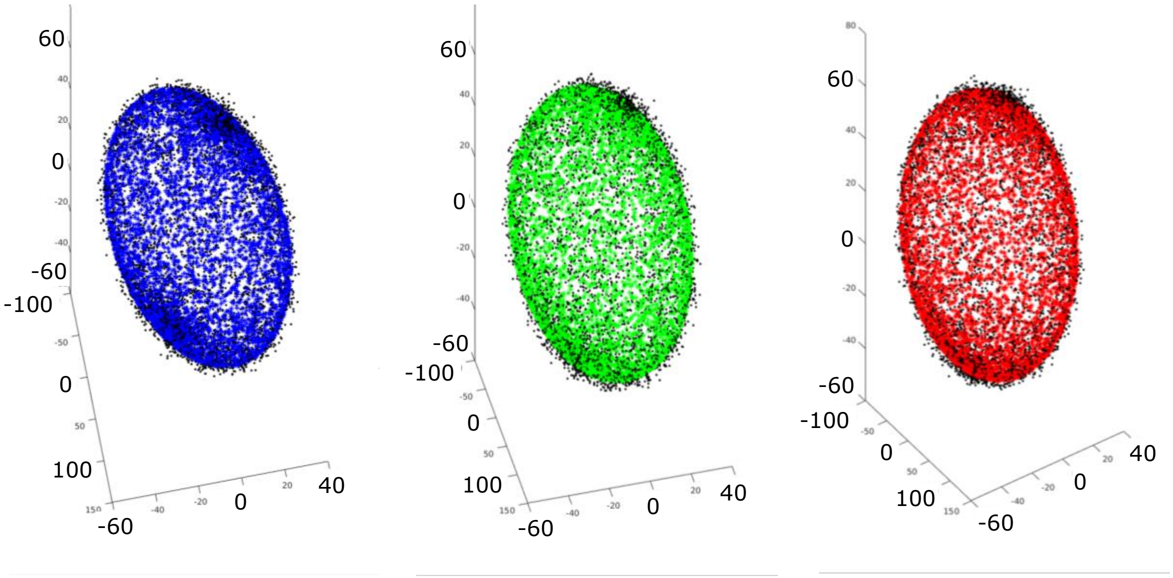


Figure 4.2: From left to right, registration of the previously described ellipsoids, ICP in blue, robust ICP in green and CPD in red.

A visual inspection can establish the convergence of the three algorithms. Besides the initial and final images, iteration by iteration movies clearly shows the monotonous convergence of all three algorithms, which can also be verified when looking at the evolution of each cost function, as presented in Figure 4.3.

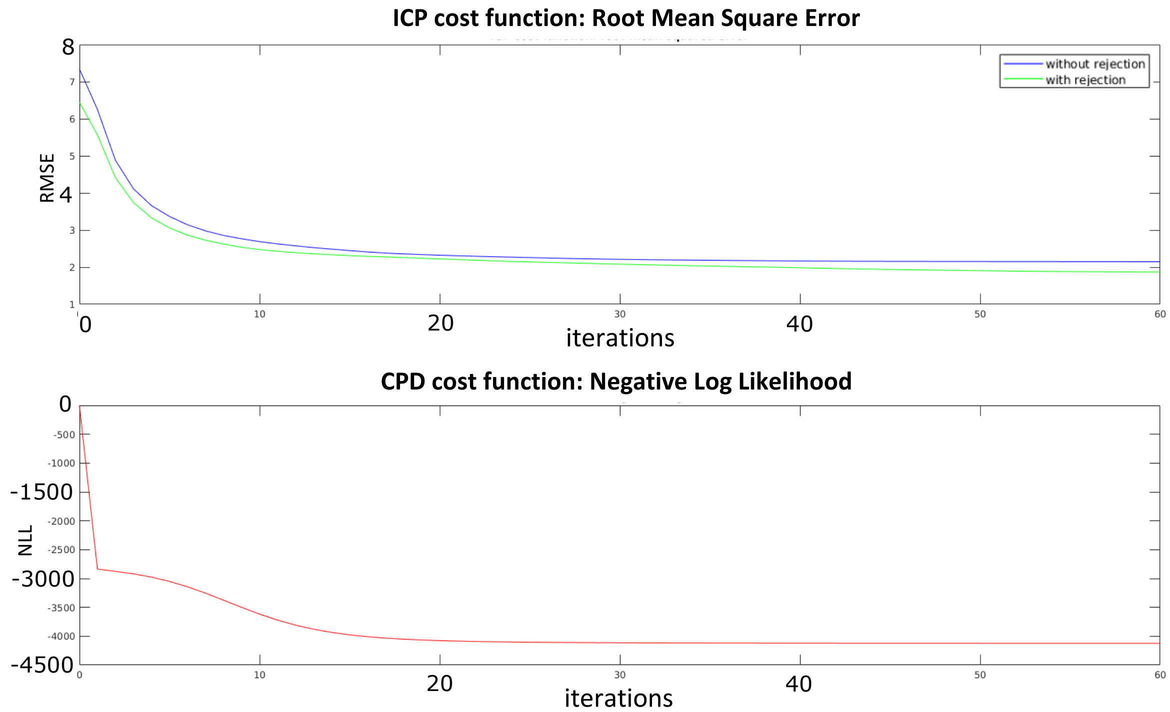


Figure 4.3: Evolution of the three cost functions for the registration of ellipsoids. Datasets of 5000 points were generated, with 2 mm noise and a rejection rate of 10% for robust ICP. Top, the RMSE, cost function of ICP and robust ICP, Bottom, the Negative LogLikelihood of the registration transformation, cost function of the CPD algorithm.

4.2. Results of Experiment 2

Figure 4.4 shows the effect of perturbation of each transformation parameter away from the gold standard on the initial Root Mean Squared Error, tested on the Sawbones lumbar phantom. Mostly, the value of the objective function is plotted for each perturbation of the registration parameters. Each of the graphs can be compared to one of the two ideal behaviors presented in Figure 3.1. The top row can be compared to an ideal rotation, and the bottom row can be compared to an ideal translation.

One detail can be noted. The error is not 0 when no transformation is made on the two datasets. This is explained by the noise added to one of the two datasets whose standard deviation is 2mm. This value of 2 mm is coherent with both the size of the diameter of a voxel in CT imaging and the localization error of the landmark using infrared tracking. There are, however, variations of this minimum, as well as the rest of the curves, which come from the randomness of the subsampling that could theoretically have some regions of the bone surface artificially depleted in points. For memory, the random subsampling is operated by choosing a predefined number of points out of the total list, independently of their position or repartition. A visual assessment can be made on the different views of Figure 3.5, but even a few hundreds of data points seem enough to correctly represent the bone surface.

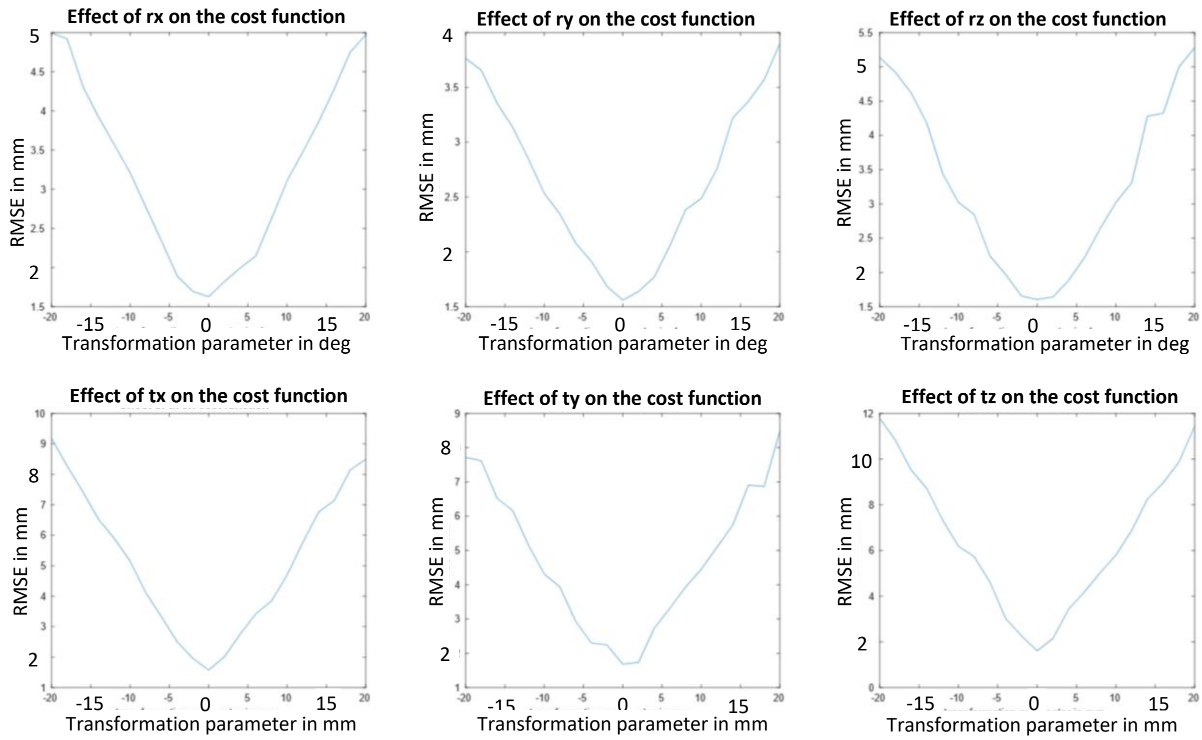


Figure 4.4: measure of the initial value of the Root Mean Squared Error for variations of the 6 transformation parameters for the Sawbones L1 phantom subsampled to 2000 points. Top to bottom and Left to Right, the modified parameter are the angle of rotation around the x-axis, the angle of rotation around the y-axis, the angle of rotation around the z-axis, the translation along the x-axis, the translation along the y-axis, the translation along the z-axis

Figure 4.5 uses the TRE to measure how well each of the registration algorithms can recover from a perturbation of the gold standard transformation for the Sawbones phantom. Each subgraph corresponds to the perturbation of one of the translation or rotation parameters. The resulting TRE values curves have random fluctuations around a constant, whereas theoretical behavior would suggest a U shape where large perturbations might result in partial recovery of the gold standard transformation. However, the three algorithms are quite robust, the values of parameters that make the TRE drastically increase are much larger than the intervals presented here, as can be seen in Figure 4.6. The rotation angle that makes the TRE increase sharply is around 70°, while the amplitude of transformation in Figure 4.5 is 20°. In this range, the TRE has a roughly constant value with random fluctuations that originate in the added noise and the different subsampled point distributions. This shows that the magnitude of the initial misregistered transformation parameter is not the limiting effect on the registration error – at least not for misregistration of the critical magnitude expected. Comparing with the minimum value of the RMSE in Figure 4.4, a TRE of 1mm seems to correspond to the minimum error of the cost function, which is coherent.

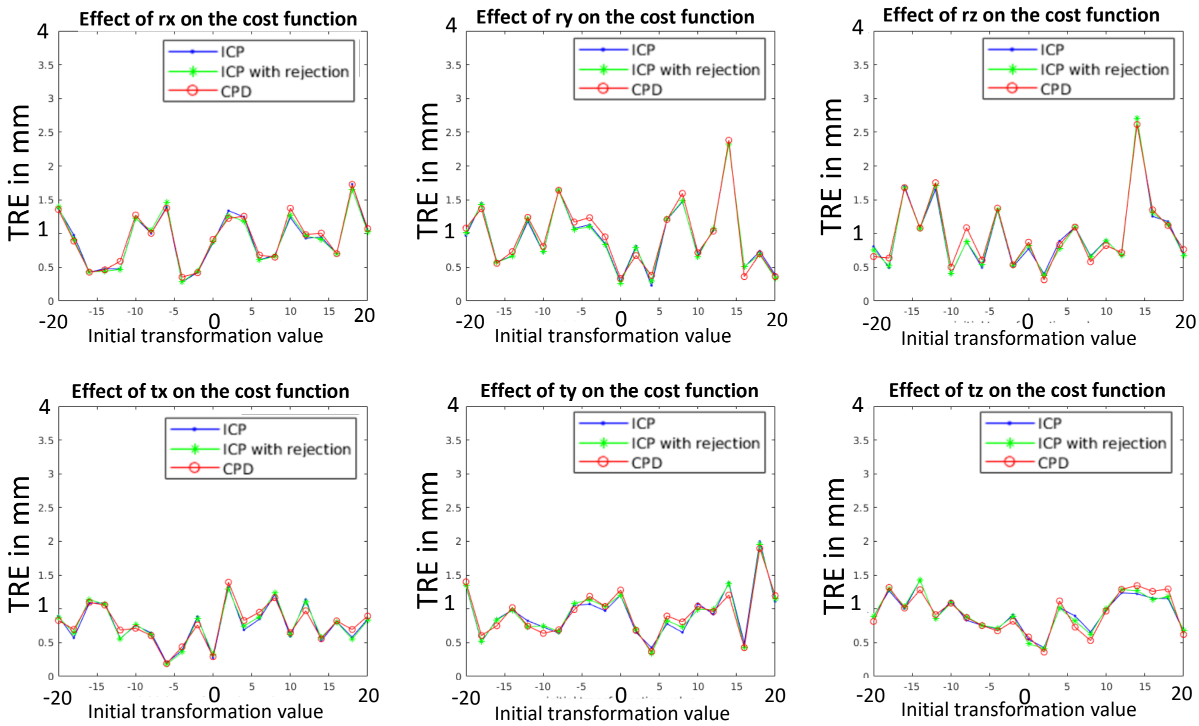


Figure 4.5: Effect of each transformation parameter on the Target Registration Error for Sawbones phantom. The top row shows the effect of rotation angles, the bottom row shows the effect of translations. The first column shows transformation around or along the x-axis, the second column the y-axis, the third column the z-axis. Blue curves refer to classic ICP, green curves to robust ICP and red curves to CPD.

Figure 4.6 and Figure 4.7 show the effect of the number of points in the two datasets on the final Target Registration Error, as well as the effect of the number of iterations for different amplitudes of transformation. The graphs show that a large amplitude of misregistration is needed to make the three algorithms diverge. Registration fails when the final TRE is not significantly smaller than the initial TRE. Visual evaluation of Figure 4.6 and Figure 4.7 show that happening for rotation angles of 70° and larger around the dorsoventral axis, and not at all for translation along the axis. Therefore, according to Figure 4.7, choosing sets of 1000 points with 40 iterations of the registration algorithms shows very stable behavior of the TRE for an extensive range of rotation angles about this axis. Mean and standard deviations of the data presented in Figure 4.8 are presented in Table 4.1. No significant difference can be visually observed between the three algorithms, which is confirmed by an ANOVA presented in Table 4.2. Observation of the top and bottom part of Figure 4.6 and Figure 4.7 show that the translation parameter is entirely stable. For all purposes investigated in the present work, this ensures the convergence of the three algorithms for the range of transformations tested, on the Sawbones phantom L1 scan. Discrepancies in the monotony of the error decrease will be addressed in the section 5.2.

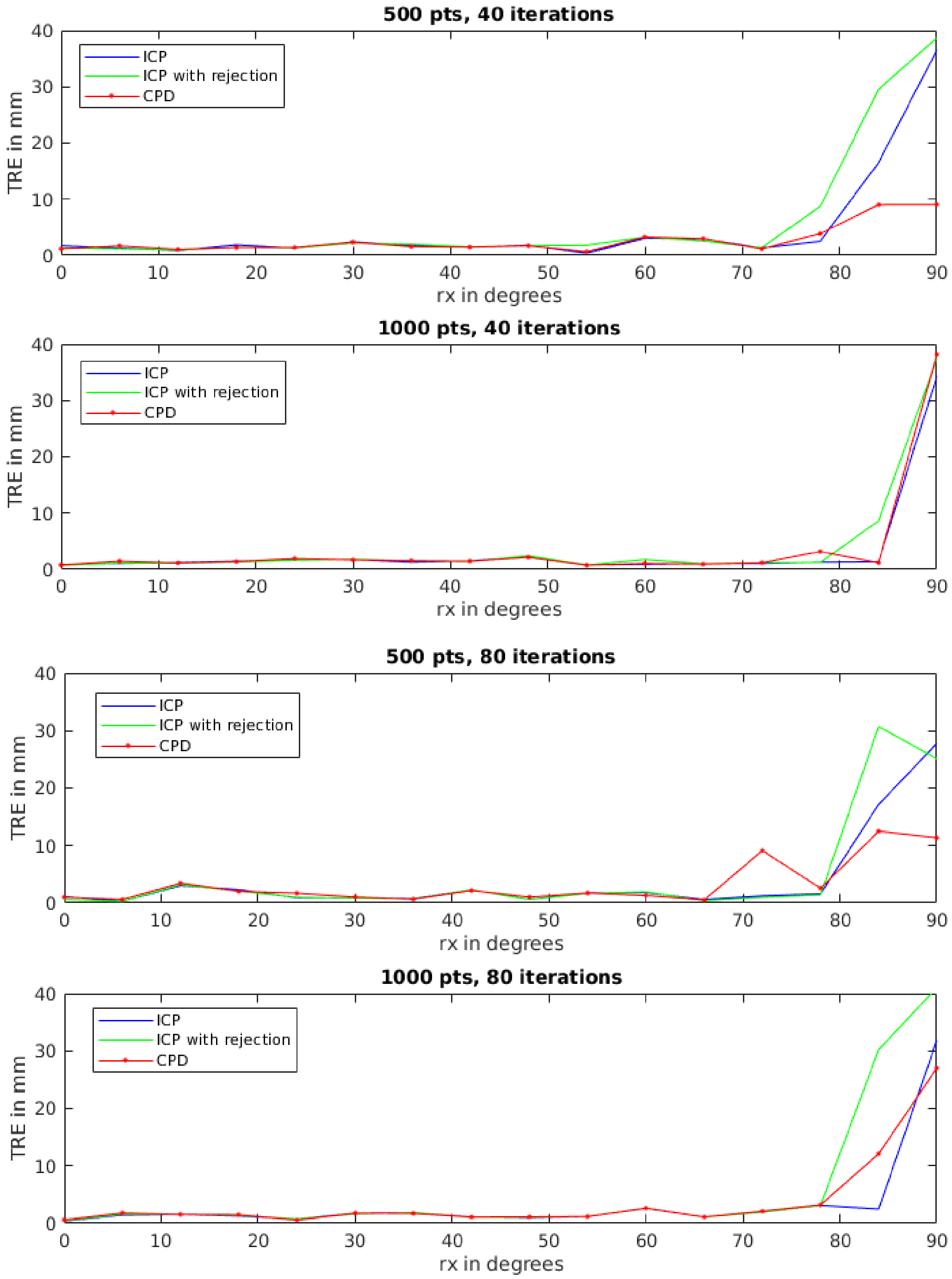


Figure 4.6: Amplitude of transformation parameter for which the registration of the Sawbones L1 phantom fails. Rejection rate = 5%, each point is an average of 5 tests. Top four, rotation around the x axis, 500 points or 1000 points subsampling, after 40 or 80 iterations.

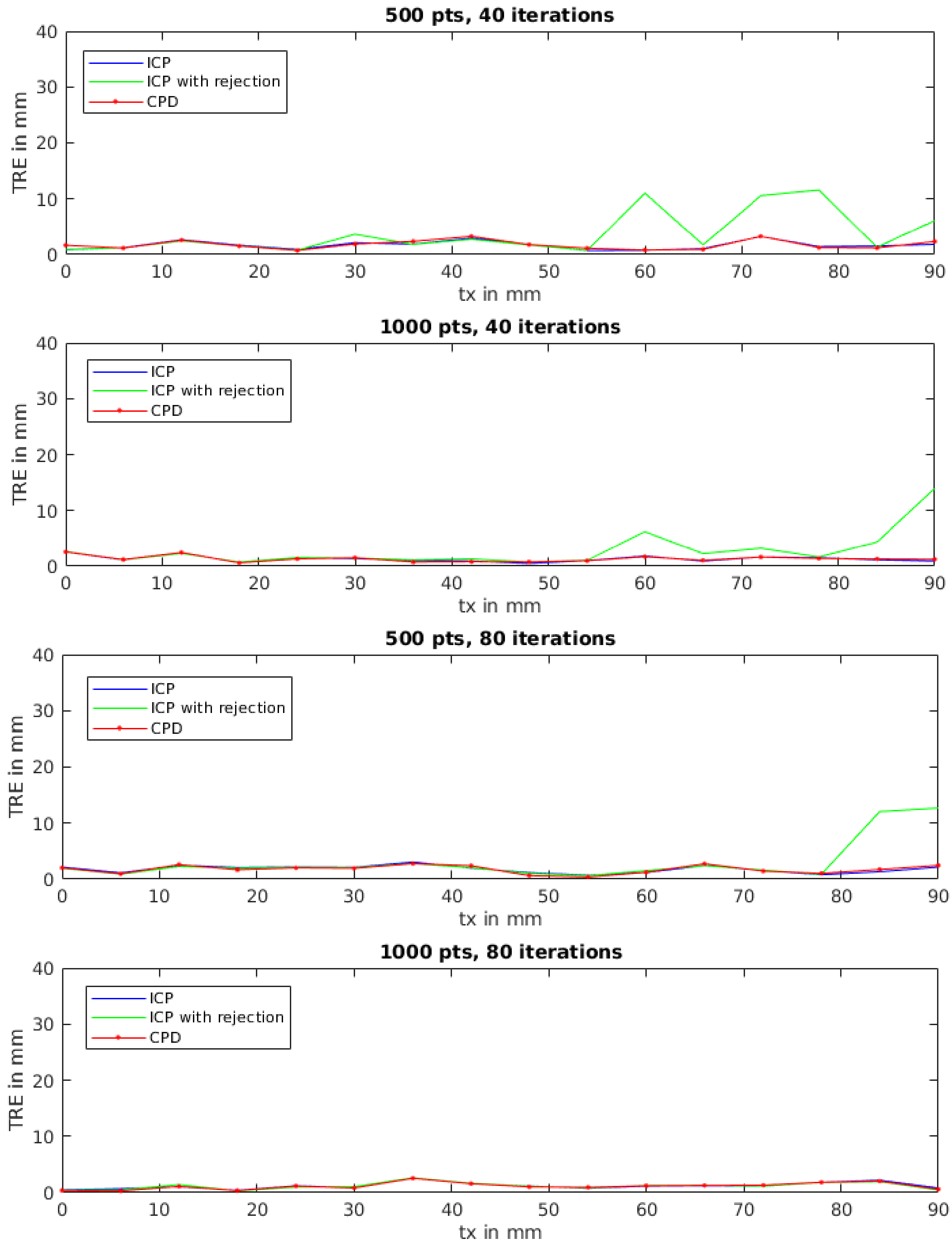


Figure 4.7: Amplitude of transformation parameter for which the registration of the Sawbones L1 phantom fails. Rejection rate = 5%, each point is an average of 5 tests. Translation along the x axis, 500 or 1000 points subsampling after 40 or 80 iterations.

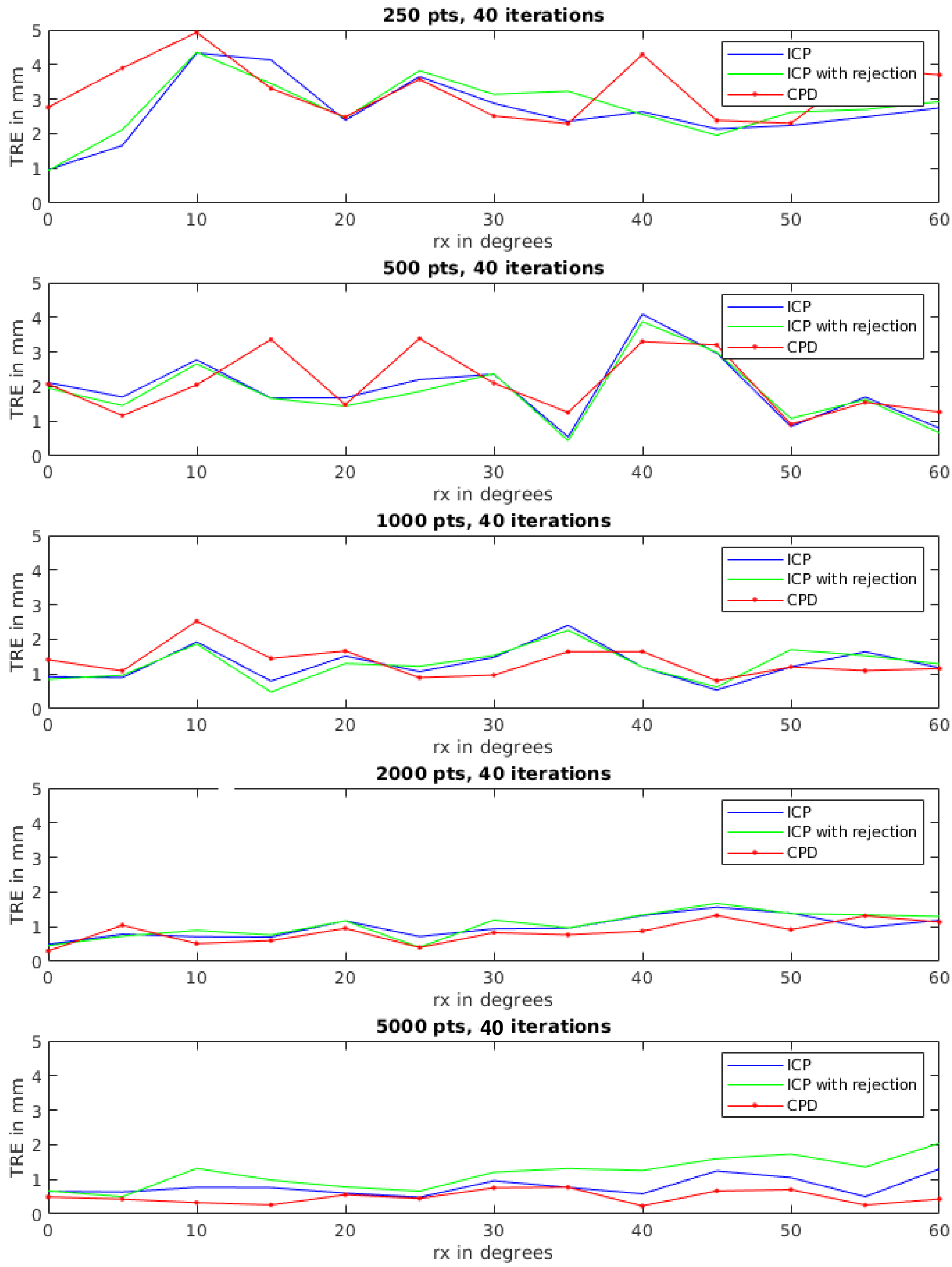


Figure 4.8: effect of the sampling on the algorithm convergence for the scan of L1 of the Sawbones phantom after 40 iterations

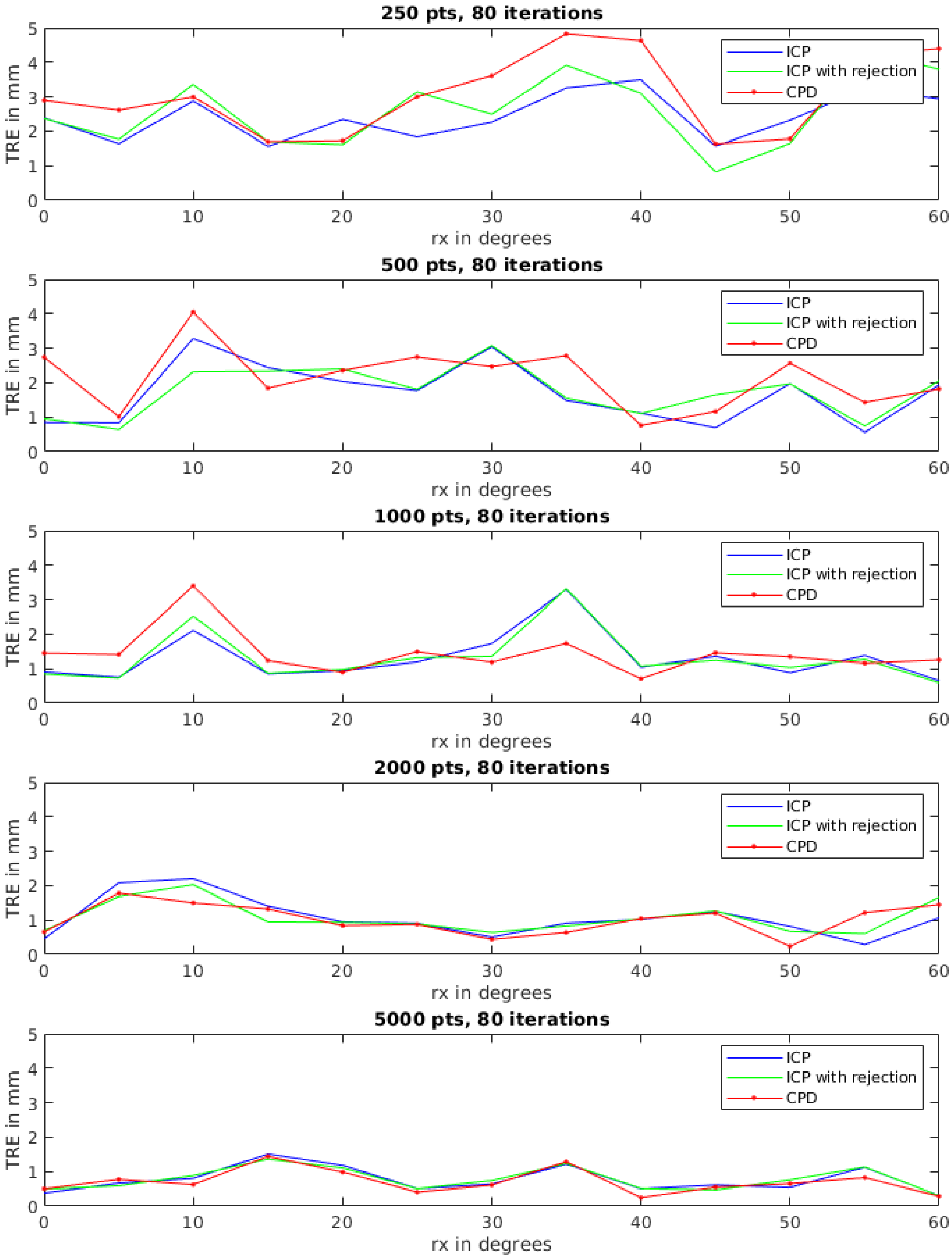


Figure 4.9: effect of the sampling on the algorithm convergence for the scan of L1 of the Sawbones phantom after 80 iterations

Table 4.1: Averages and standard deviations of the variation of subsampling and number of iteration on the final TRE for L1 of the Sawbones phantom as presented in Figure 4.8 and ***Error! Reference source not found.***

Iterations	Number of points	Mean and std of TRE (mm)		
		ICP	robustICP	CPD
40	250	2.66 ± 0.9	2.79 ± 0.9	3.26 ± 0.9
	500	1.96 ± 1.0	1.85 ± 0.9	2.08 ± 0.9
	1000	1.29 ± 0.5	1.30 ± 0.5	1.35 ± 0.5
	2000	0.99 ± 0.3	1.05 ± 0.4	0.84 ± 0.3
	5000	0.80 ± 0.3	1.19 ± 0.5	0.49 ± 0.2
80	250	2.44 ± 0.7	2.62 ± 1.1	3.08 ± 1.2
	500	1.69 ± 0.9	1.74 ± 0.7	2.13 ± 1.0
	1000	1.32 ± 0.7	1.32 ± 0.8	1.44 ± 0.6
	2000	1.07 ±0.6	1.07 ± 0.5	1.02 ± 0.5
	5000	0.77 ± 0.4	0.78 ± 0.3	0.71 ± 0.4

A three-factor ANOVA was conducted on this data with a threshold of $\alpha=0.05$. The calculation was run through the SPSS software. The results of the ANOVA are presented in the Table 4.2.

Table 4.2: Three-way ANOVA analysis of the results presented in the Figure 4.8 to ***Error! Reference source not found.***

Source	Sum of Square	Degree of freedom	Mean Square	F ratio	P value
Algorithm	0.501	2	0.25	0.522	NS
Number of Points	78.874	4	19.718	41.155	P<0.0001
Iterations	0.083	1	0.083	0.173	NS
Algorithm and Number of Points	3.038	8	0.38	0.793	NS
Iterations and Algorithm	0.272	2	0.136	0.284	NS
Iterations and Number of Points	0.398	4	0.099	0.208	NS
All 3 factors	0.39	8	0.049	0.102	NS
Error	57.495	120	0.479		

A post-hoc Tukey HSD test was carried on the main effect of the Number of Points, showing significant difference for each comparison except between 1000 and 2000 points, and between 2000 and 5000 points. This confirms the absence of significant statistical difference between the three algorithms and the two number of iterations.

Figure 4.10 shows the effect of rejecting outliers at each iteration of the registration algorithms, for datasets of 2000 points iterated 40 times. Each plot shows the result of 500 tests. The distributions show very close average values of the TRE independently of the fraction of rejected points, although a slight increase happens for a rate of 20%. This rejection rate is the trade-off between the relative importance of the noise and outliers in the robust ICP registration process. This indicates that a noise amplitude of 2mm has a more substantial effect than outliers on the registration accuracy for this dataset.

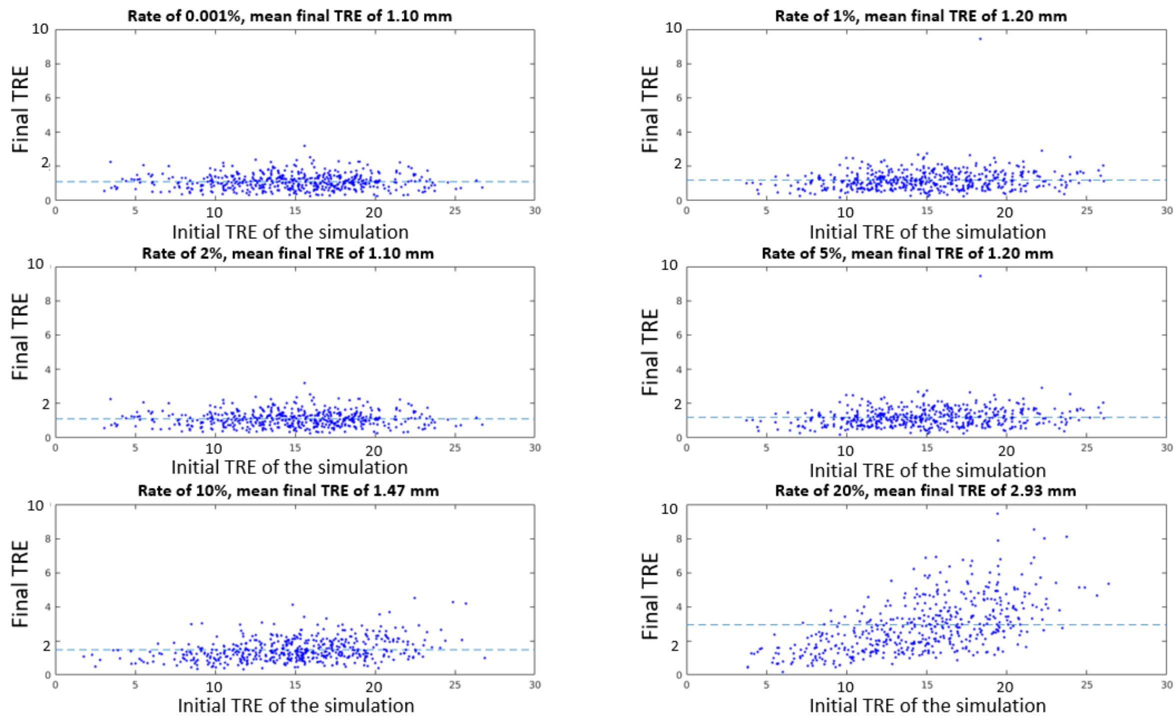


Figure 4.10: effect of the rate of rejection on the final TRE distribution of robust ICP on L1 Sawbones phantom. The dotted line represents the average value of the scatter plot

Figure 4.11 shows the result of 500 registration tests carried on datasets of 2000 points with a maximum of 40 iterations and a rejection rate of 1%. Each point corresponds to one registration test, where the abscissa is the initial TRE value, and the ordinate is the final TRE. A distribution of final TRE value for each technique can be approximated using this data. One can see that for a broad range in initial TRE values (25mm would be more significant than one would expect in the clinic), the three registration techniques yield sub-millimeter registration accuracy.

One can see that all three algorithms appear to yield stable results across a broad range of initial TRE values, as can be seen by comparing the width of the distributions on the right of Figure 4.11. The three distributions have a width between 0.3 and 0.36mm, making the final TRE well below the required resolution, with only the tail peaking above 2mm.

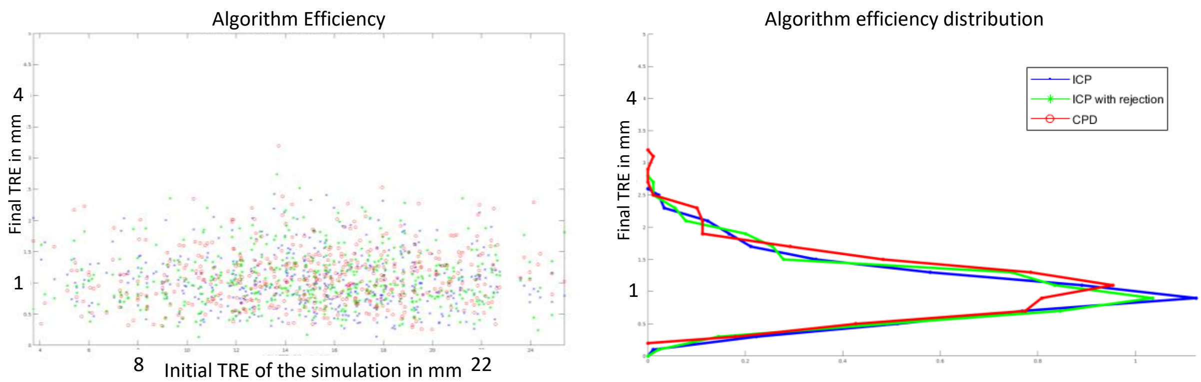


Figure 4.11: From left to right, values of the final TRE for 750 tests, on datasets of 2000 points, with 60 iterations given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green corresponds to robust ICP and red to CPD.

4.3. Results of Experiment 3

Just as in Figure 4.8 with the Sawbones phantom, Figure 4.12 shows the effect of the subsampling fraction and the maximal number of iterations on the convergence of the three algorithms for the registration of ultrasound and CT scans of vertebra L3 for different amounts of initial misregistration. The first notable point is the minimum value around 10mm of the TRE – much larger than the ~1 mm TRE obtained with the Sawbones phantom. This corresponds to a root mean squared error on the target points of around 1cm. The possible causes of this very high value will be investigated and detailed in the next chapter. Subsampling shows many more variations in the final value of the TRE than above. Indeed, if for the Sawbones phantom, the final TRE curves were quasi flat for transformation parameter below 70°, for porcine L3 cadaver, increases appear clearly at 30° and greater. The average value of each curve, as well as standard deviation are presented Table 4.3.

No visual difference could be extracted from the curves corresponding to 1000 points, 2000 points, and 5000 points, whether they compute 40 or 80 iterations. Because each case was made with different sets of points and most of the tests below are made on the same subsampling of the datasets with variations of the transformation parameters, a number of points of 2000 points and 40 iterations were chosen as a precaution, in order to mitigate the potential number of local minima.

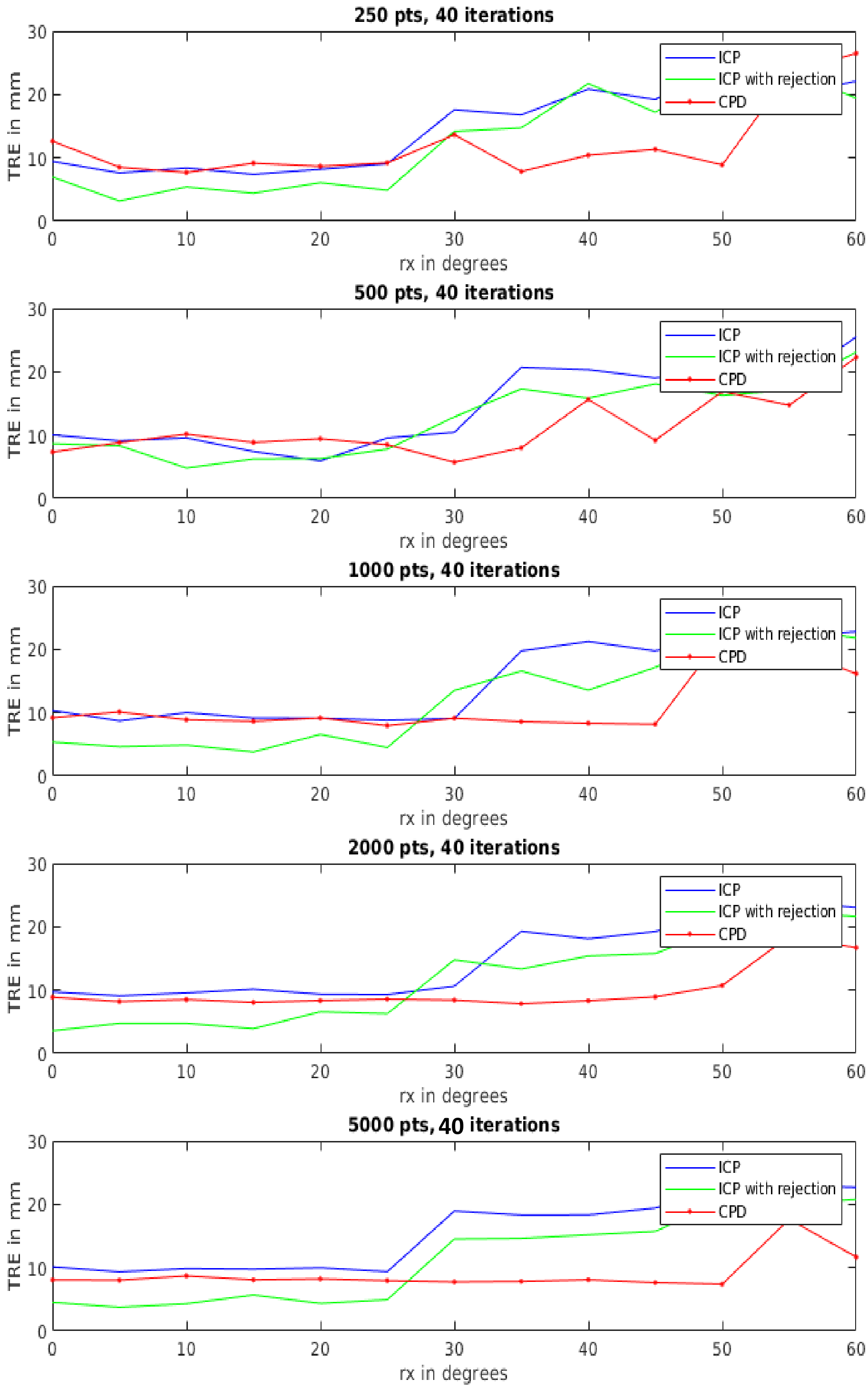


Figure 4.12: Effect of the number of points on the convergence of algorithms on the scans of L3 porcine phantoms for maximum 40 iterations

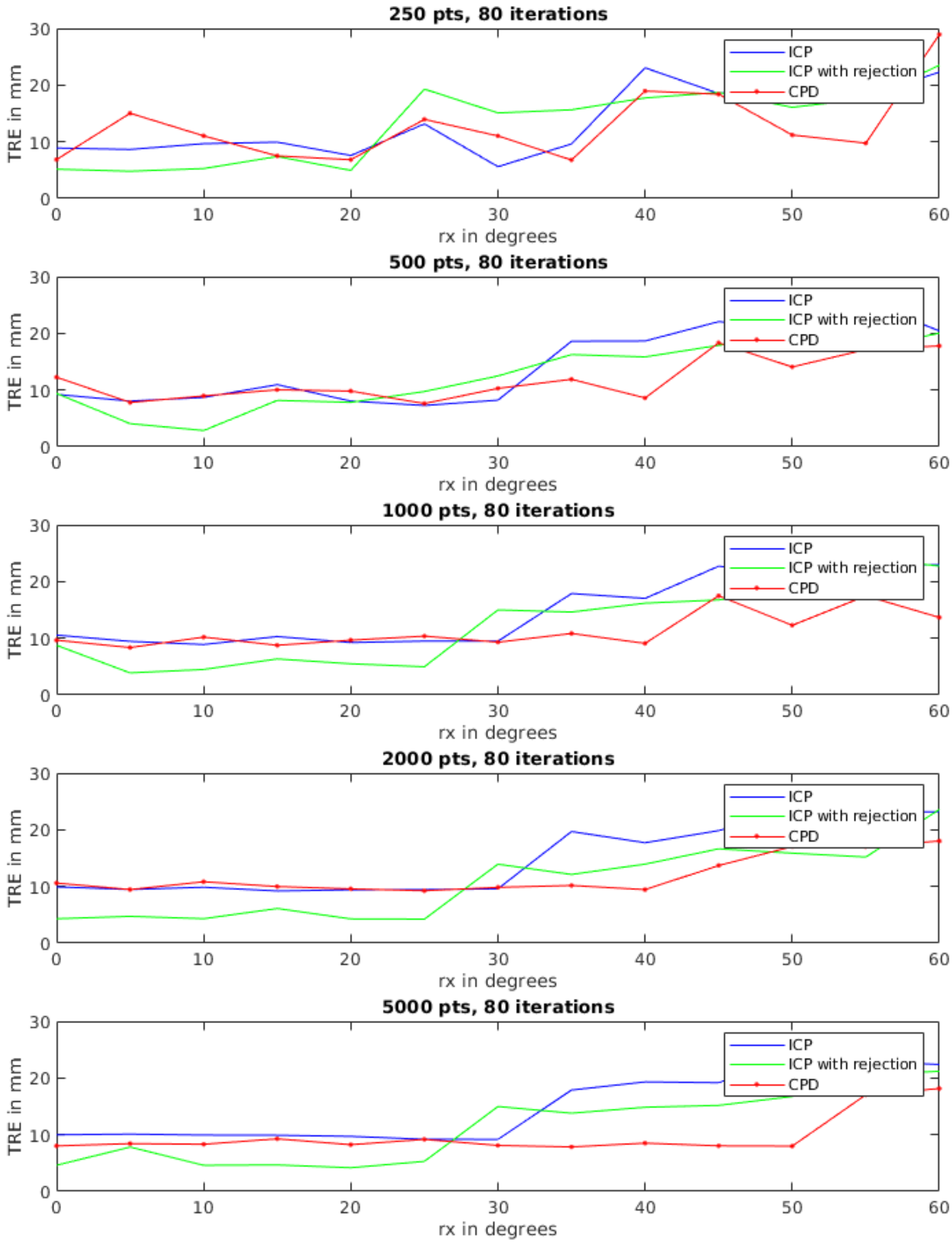


Figure 4.13 : Effect of the number of points on the convergence of algorithms on the scans of L3 porcine phantoms for maximum 80 iterations

*Table 4.3: Averages and standard deviations of the variation of subsampling and number of iteration on the final TRE for L3 of porcine cadaver as presented in **Error! Reference source not found.**2 and 4.13.*

Iterations	Number of points	Mean TRE±STD (mm)		
		ICP	robust ICP	CPD
40	250	14.75 ± 6.5	12.72 ± 7.9	12.14 ± 6.0
	500	14.34 ± 6.5	15.48 ± 5.8	14.79 ± 6.1
	1000	12.54 ± 5.8	11.45 ± 6.9	12.45 ± 7.1
	2000	11.21 ± 4.7	9.00 ± 2.8	11.32 ± 3.1
	5000	14.87 ± 6.3	13.53 ± 6.1	14.93 ± 6.2
80	250	12.14 ± 7.5	13.17 ± 6.6	10.71 ± 6.4
	500	11.27 ± 4.8	12.79 ± 6.4	11.92 ± 3.3
	1000	14.92 ± 6.1	14.38 ± 6.7	14.90 ± 6.0
	2000	11.77 ± 7.1	12.48 ± 5.9	11.42 ± 6.9
	5000	10.01 ± 3.5	11.89 ± 3.8	9.78 ± 3.5

Once again, a three-factor ANOVA was computed using SPSS. No significant effect was discovered with a threshold of α=0.05.

Figure 4.14 shows the effect of different rejection rates on robust ICP registration of datasets of 2000 points over 80 iterations, for registration of ultrasound and CT data of the vertebra L3. Figure 4.13 and Table 4.3 show that there is very little difference between 40 and 80 iterations for numbers of points superior to 1000, therefore selecting a maximum number of iterations of 80 instead of 40 does not have any effect on the final result. The number of local minima making plateaus of values for the final TRE, hiding the potential decrease, would prevent any relevant analysis of the variation of the rejection rate if the same subsampling of 2000 points was chosen for every 500 tests of each graph. That is why, for this graph, the subsampling was changed, taking a new set of 2000 points. Both the graphs and the average TRE values show that a rejection rate of 20% seems to minimize the final TRE.

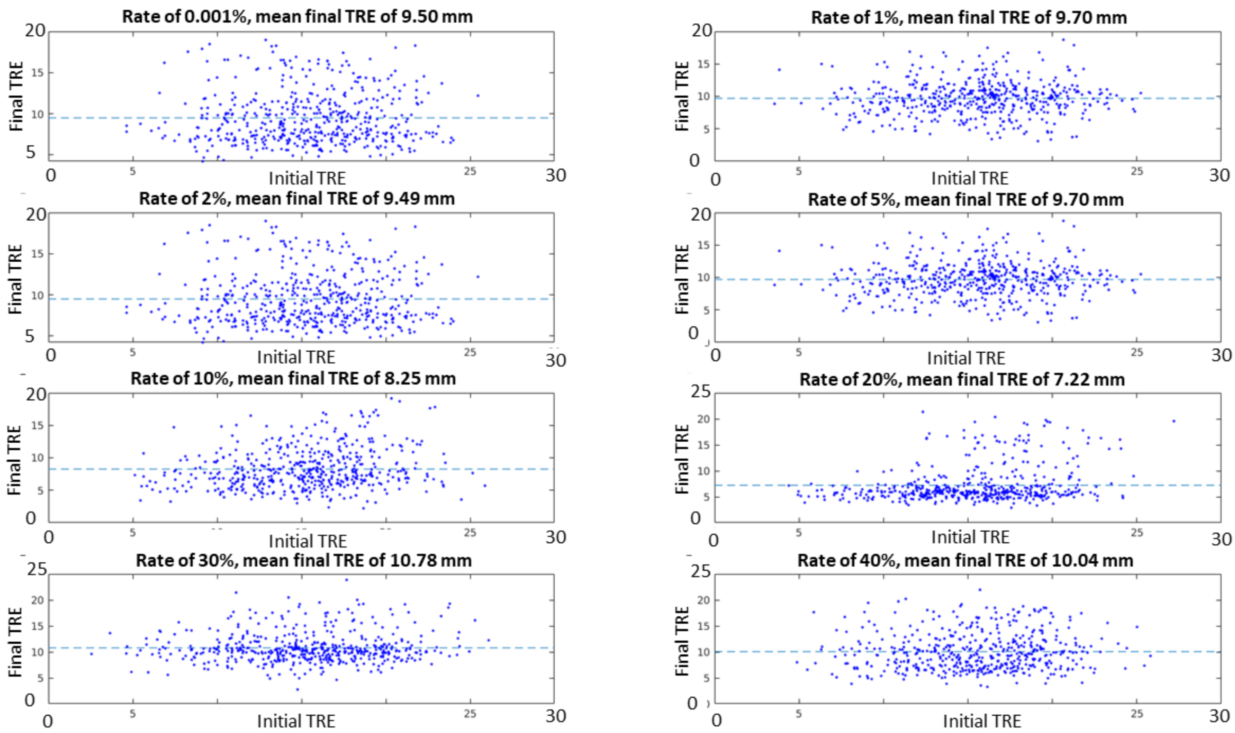


Figure 4.14: Effect of the rate of outliers rejection for robust ICP of ultrasound and CT data

Figure 4.15 to Figure 4.20 shows the effect of single transformation parameters perturbations from the gold standard alignments on the Root Mean Squared Error objective function, for vertebrae L1 to L6 with sets of 2000 points, over 40 iterations. The rejection rate used for robust ICP is 1%. When comparing these to Figure 4.4, several changes of behavior must be noted. First, the theoretical smooth 'U' shape of the objective function is not discernable for some tests. The rotational parameters harbor many more variations and of a bigger amplitude than with the Sawbones phantom, indicating multiple local minima. As the RMSE is the cost function optimized by the two ICP algorithms studied here, the registration algorithm may be lost in a local minimum. The translation parameters appear much smoother with less local minima; however, their minimum value, the zero abscissa, is not always attained for an absence of transformation. This would mean that the minimum of the cost function does not coincide with the gold transformation. This possibility will be investigated further in the next chapter. Finally, it is to be noted that the minimum value of each of the concave curves is significantly larger than for earlier experiments with the Sawbones phantom. One possible explanation would be that the noise amplitude in the data is more extensive than expected.

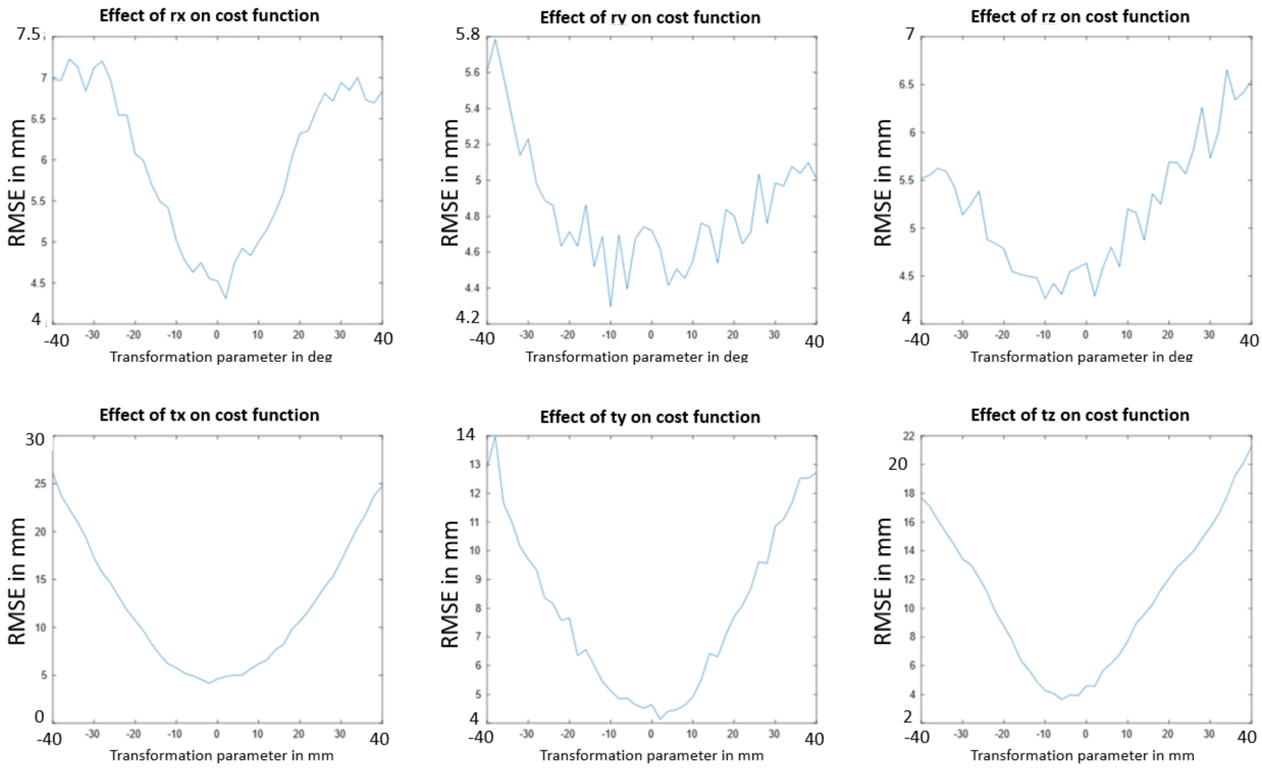


Figure 4.15: effect of each transformation parameter on the Root Mean Squared Error for registration of CT and ultrasound data of L1 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis

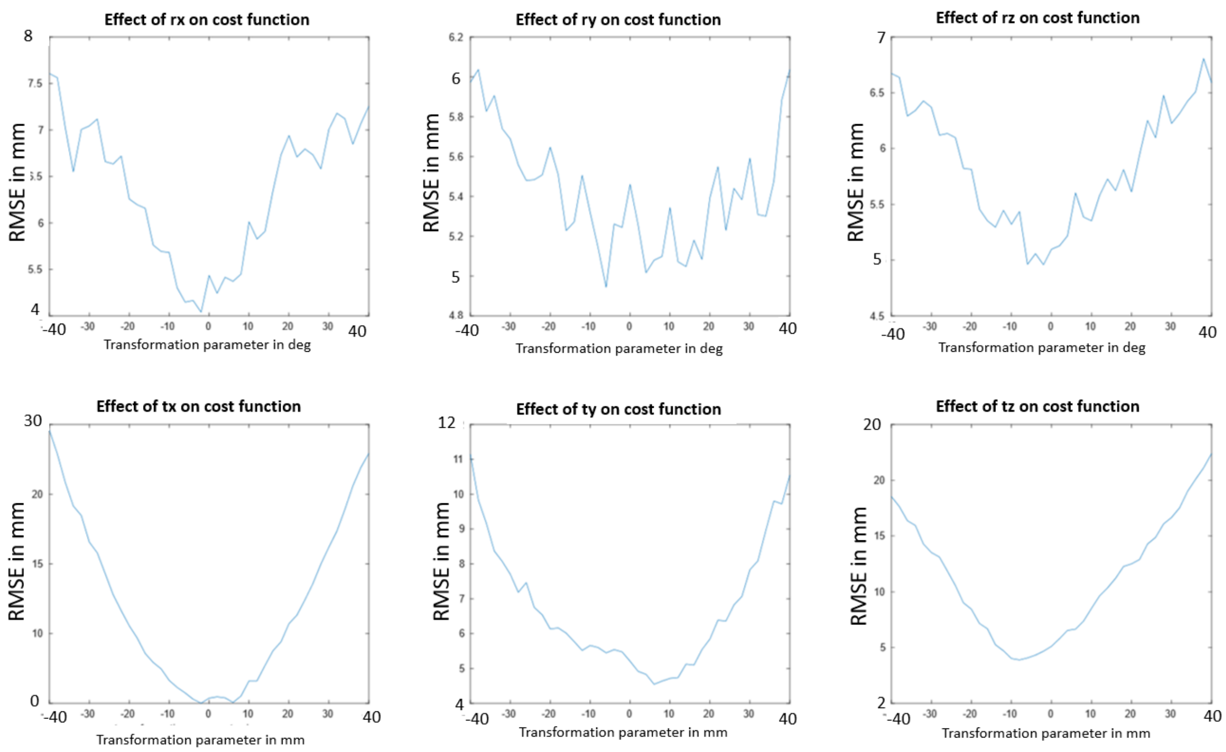


Figure 4.16: effect of each transformation parameter on the Root Mean Squared Error for registration of CT and ultrasound data of L2 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis

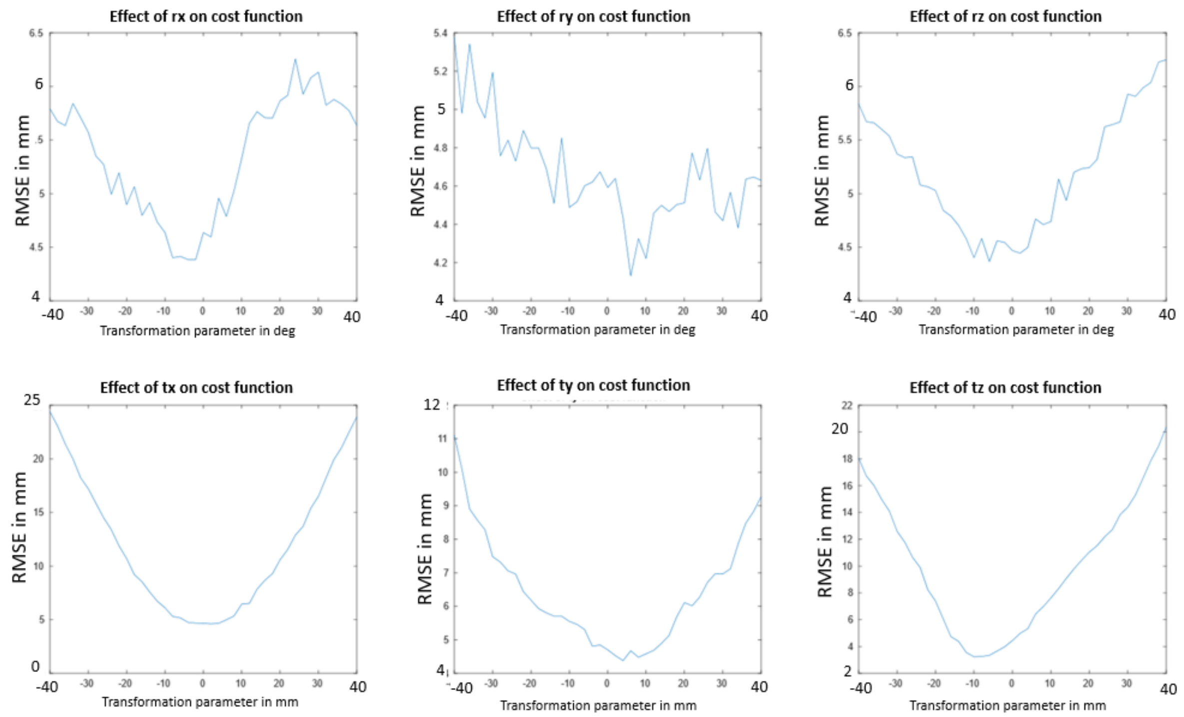


Figure 4.17: effect of each transformation parameter on the Root Mean Squared Error for registration of CT and ultrasound data of L3 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis

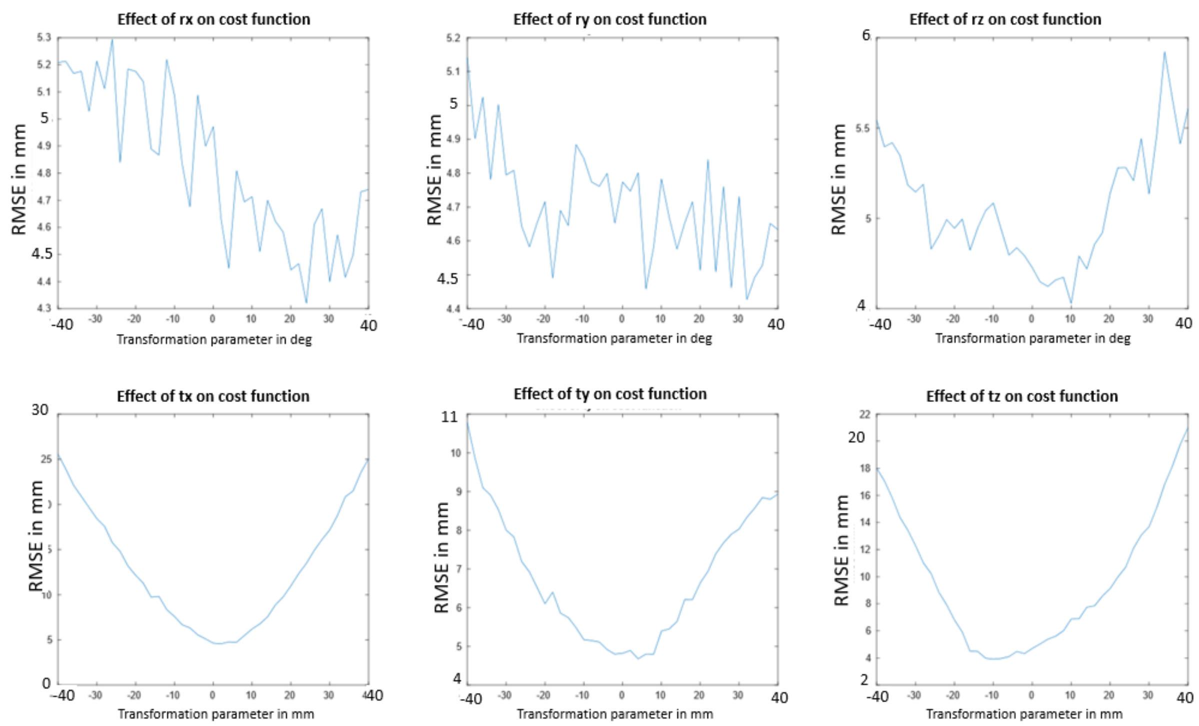


Figure 4.18: effect of each transformation parameter on the Root Mean Squared Error for registration of CT and ultrasound data of L4 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis.

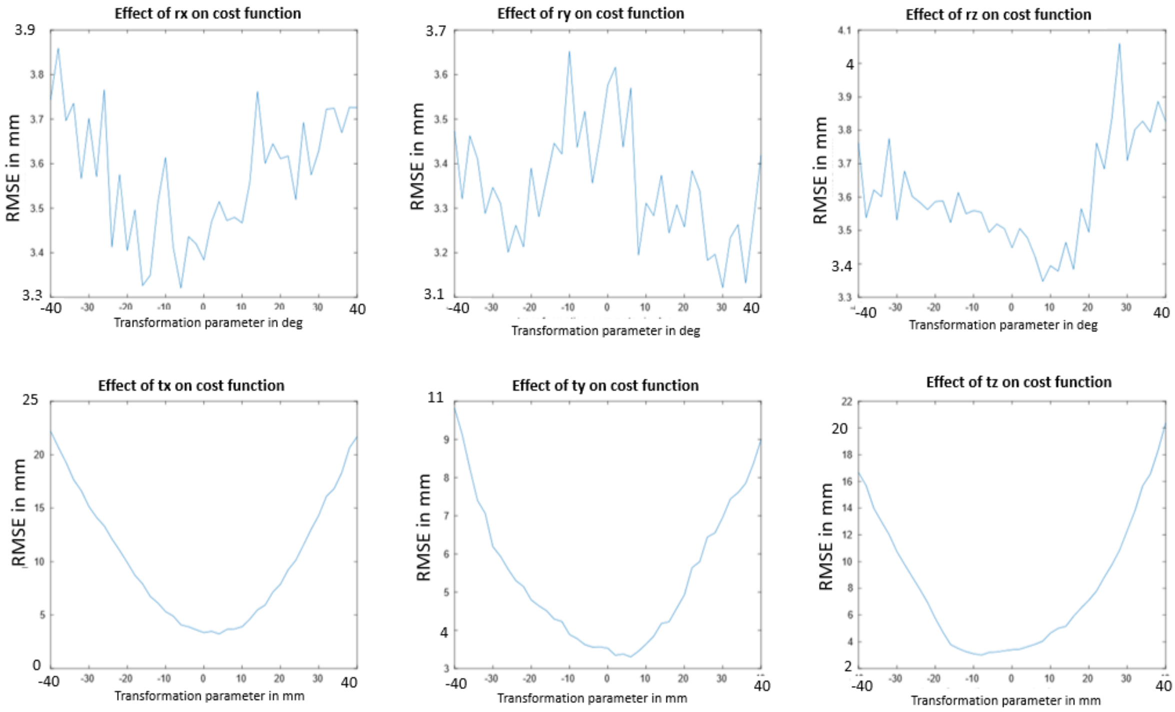


Figure 4.19: effect of each transformation parameter on the Root Mean Squared Error for registration of CT and ultrasound data of L5 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis

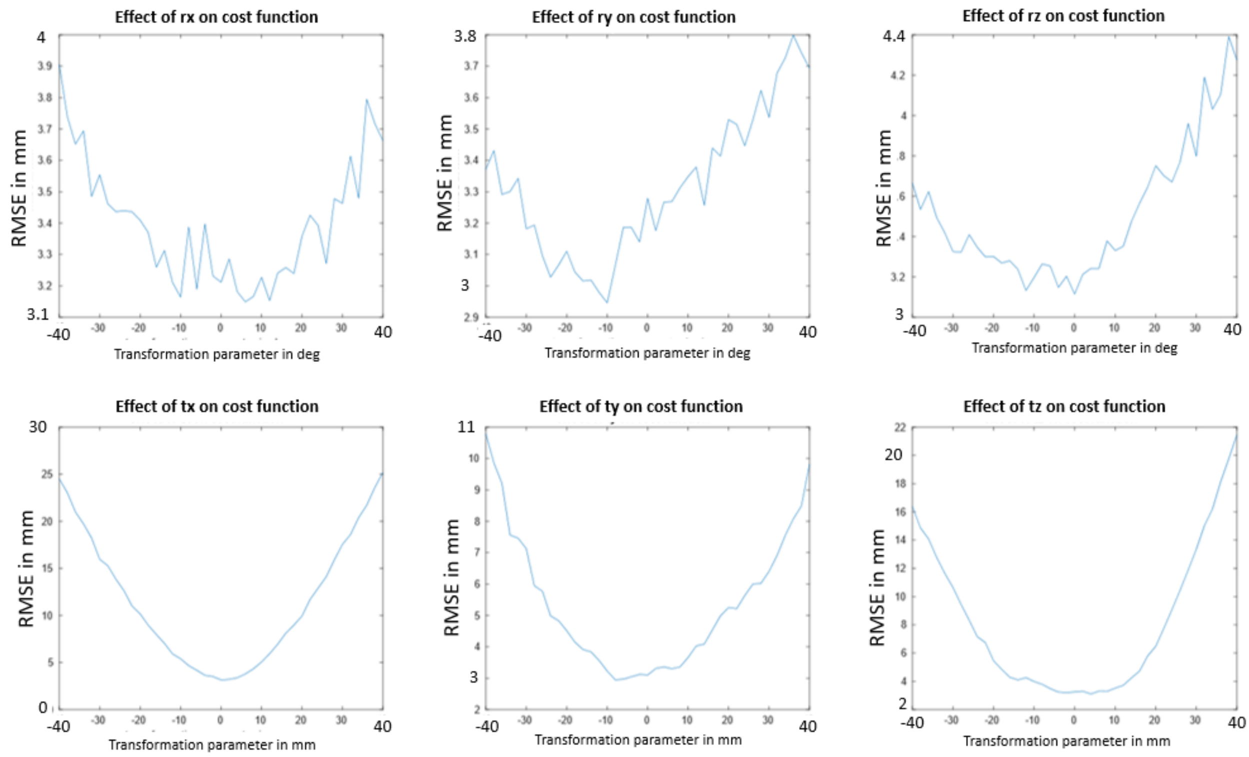


Figure 4.20: effect of each transformation parameter on the Root Mean Squared Error for registration of CT and ultrasound data of L6 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis

Figure 4.21 to Figure 4.26 below show the value of final TRE after registration from a starting point with perturbations of single transformation parameters for each vertebra L1 to L6, using the three algorithms with a limit of 40 iterations, for datasets of 2000 points and a rejection rate of 1% for robust ICP. The TRE graphs for L1, L2, L3, and L6 have minimal curvature, suggesting that the algorithms converged to their lowest level and are robust over a broad range of parameter perturbations. The minimum values of TRE oscillate just below 10 mm, which is not ideal but coherent with previous tests. These results question the accuracy of the algorithms used, but their robustness and precision are quite strong. For four out of the six vertebrae, the amplitude of the transformation parameters has minimal impact on the final TRE.

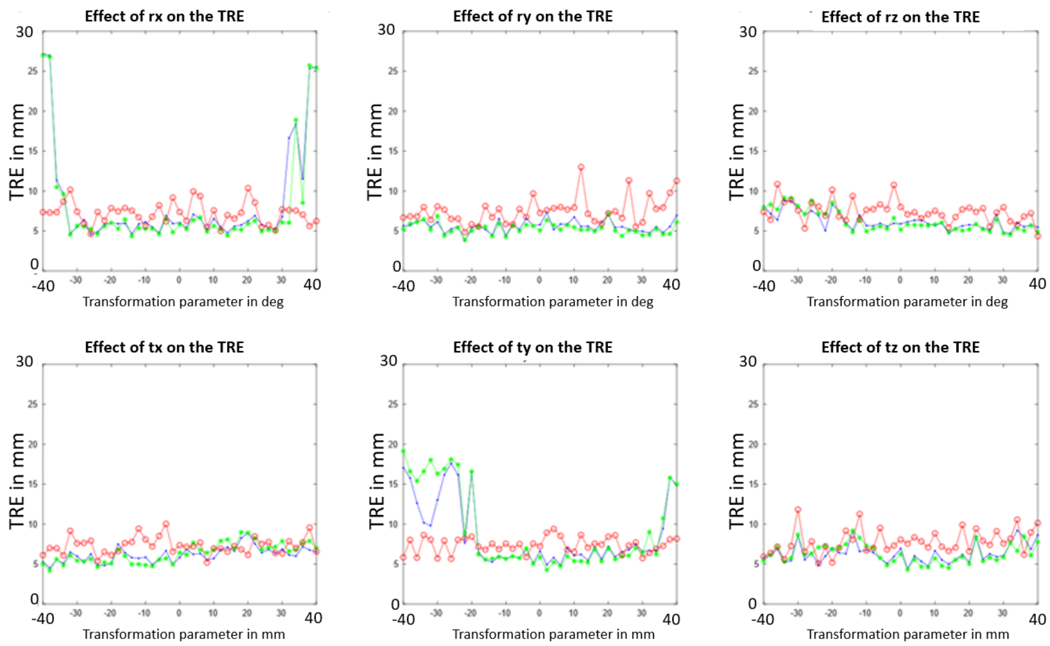


Figure 4.21: effect of each transformation parameter on the final Target Registration Error for registration of CT and ultrasound data of L1 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis. Blue curves show results for the classic ICP algorithm, green for the robust ICP and red for CPD.

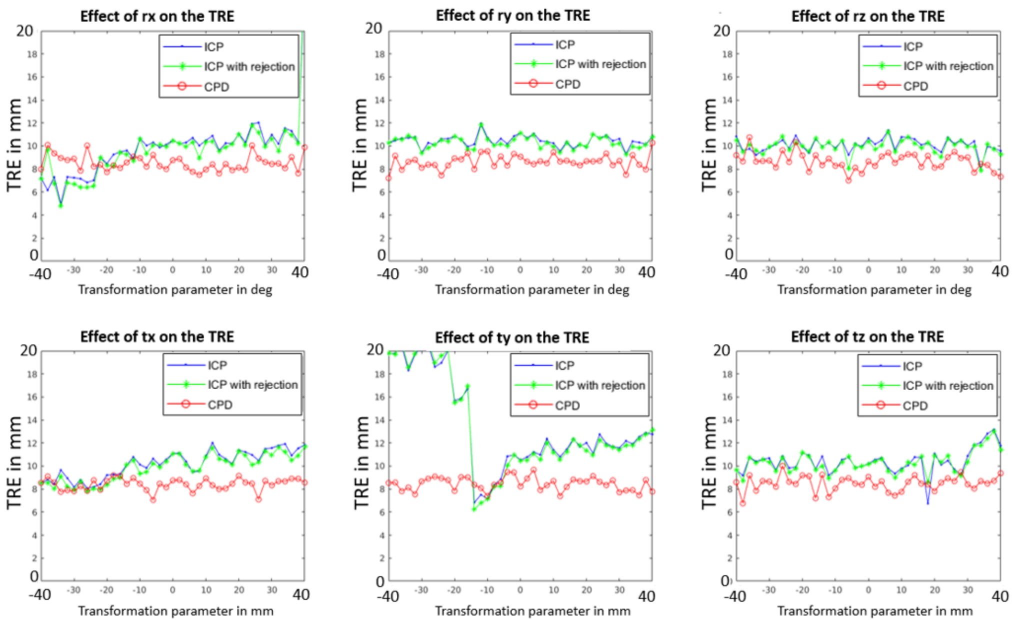


Figure 4.22: effect of each transformation parameter on the final Target Registration Error for registration of CT and ultrasound data of L2 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis. Blue curves show results for the classic ICP algorithm, green for the robust ICP and red for CPD.

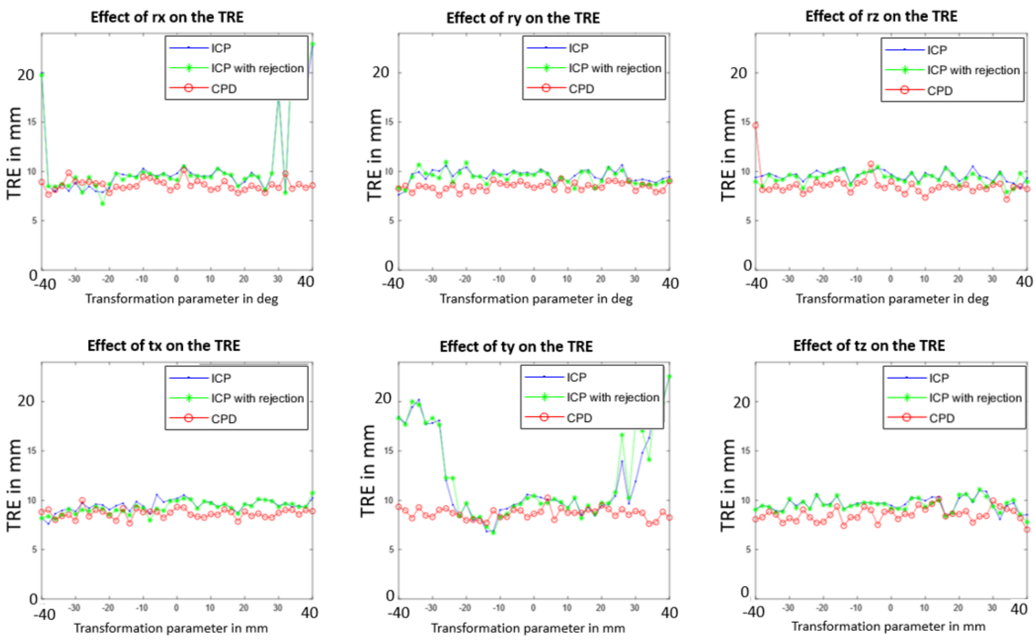


Figure 4.23: effect of each transformation parameter on the final Target Registration Error for registration of CT and ultrasound data of L3 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis. Blue curves show results for the classic ICP algorithm, green for the robust ICP and red for CPD.

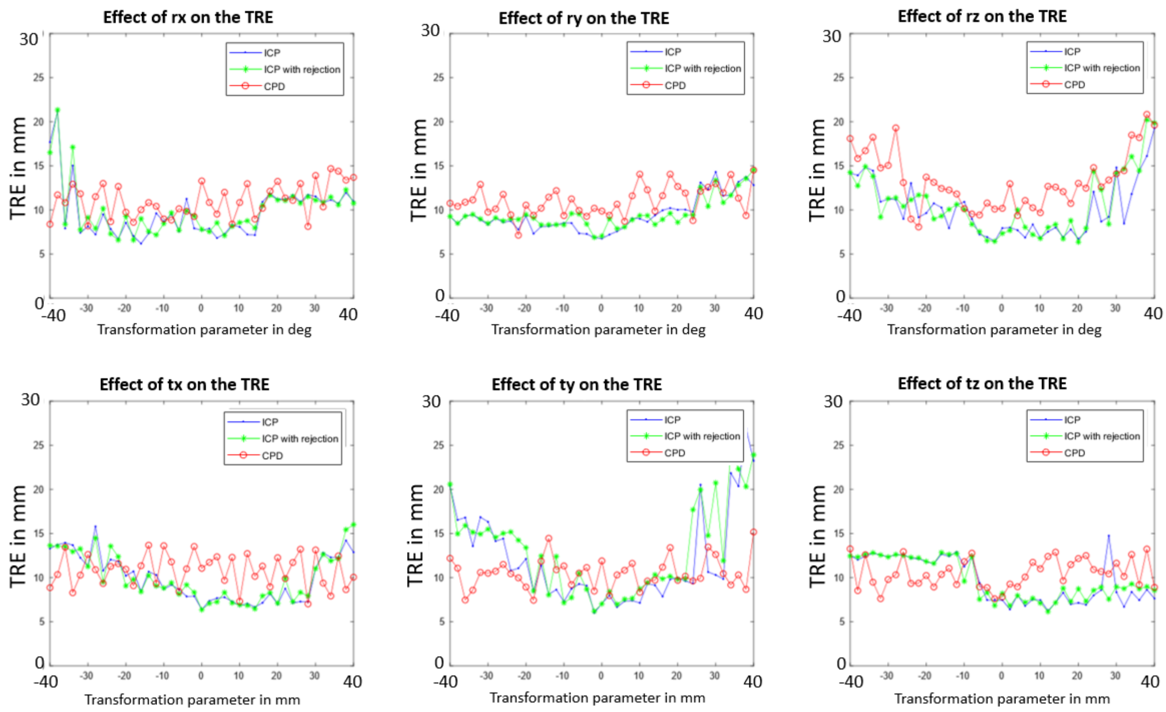


Figure 4.24: effect of each transformation parameter on the final Target Registration Error for registration of CT and ultrasound data of L4 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis. Blue curves show results for the classic ICP algorithm, green for the robust ICP and red for CPD.

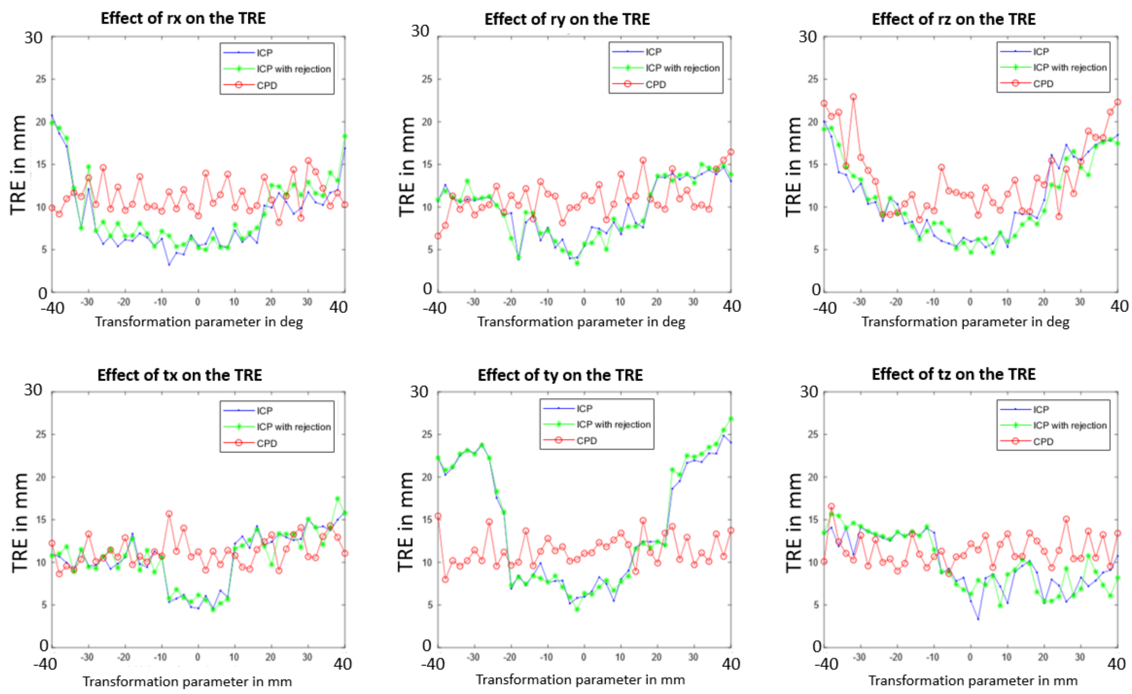


Figure 4.25: effect of each transformation parameter on the final Target Registration Error for registration of CT and ultrasound data of L5 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis. Blue curves show results for the classic ICP algorithm, green for the robust ICP and red for CPD.

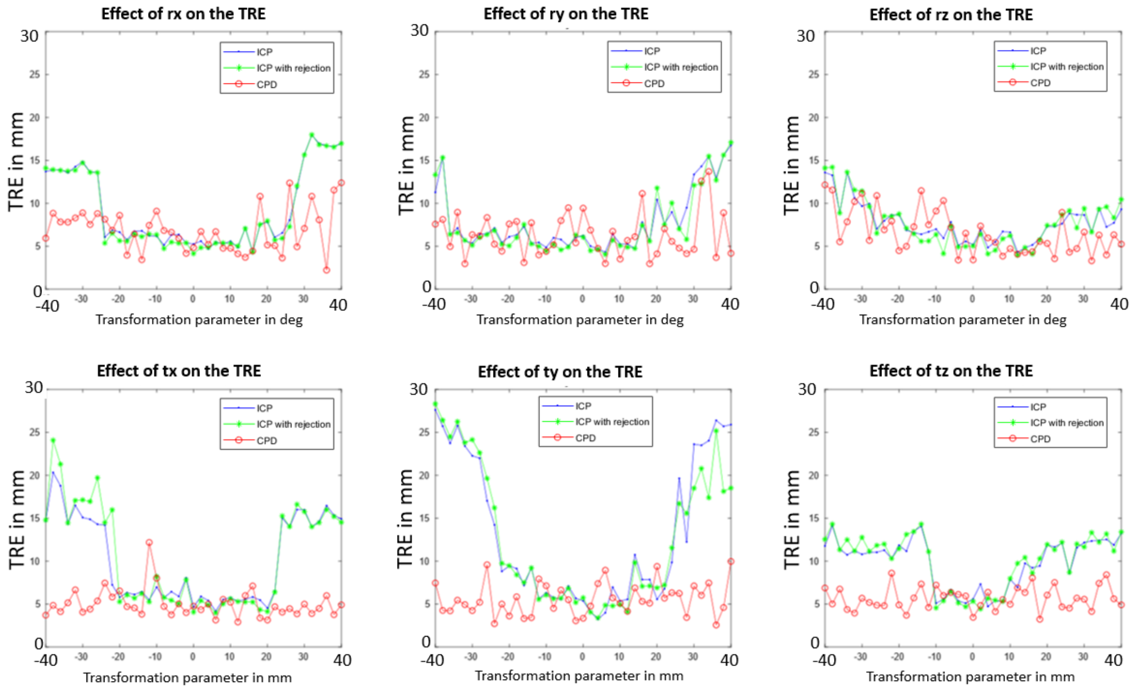


Figure 4.26: effect of each transformation parameter on the final Target Registration Error for registration of CT and ultrasound data of L6 for porcine cadavers. Top shows rotation angles, bottom shows translations. First column is the x-axis, second is the y-axis, third is the z-axis. Blue curves show results for the classic ICP algorithm, green for the robust ICP and red for CPD.

Figure 4.27 to Figure 4.32 show the results of 500 random CT-US registration tests on datasets of 2000 points over 40 iterations for each vertebra L1 to L6 of the porcine phantom. As before, the straight horizontal lines illustrate the presence of local minima in the cost functions of the algorithms, which explains the amplitude of the TRE. This comes from the fact that the same subsampling was used for every test of each graph. The distributions, however show the spread that each algorithm has at its output, i.e., the robustness of the registration. Those figures show the robustness of the CPD algorithm unambiguously. The values of TRE are, however, substantial compared to the clinical goal, which suggests a sharp discrepancy between the gold standard and the registration algorithms.

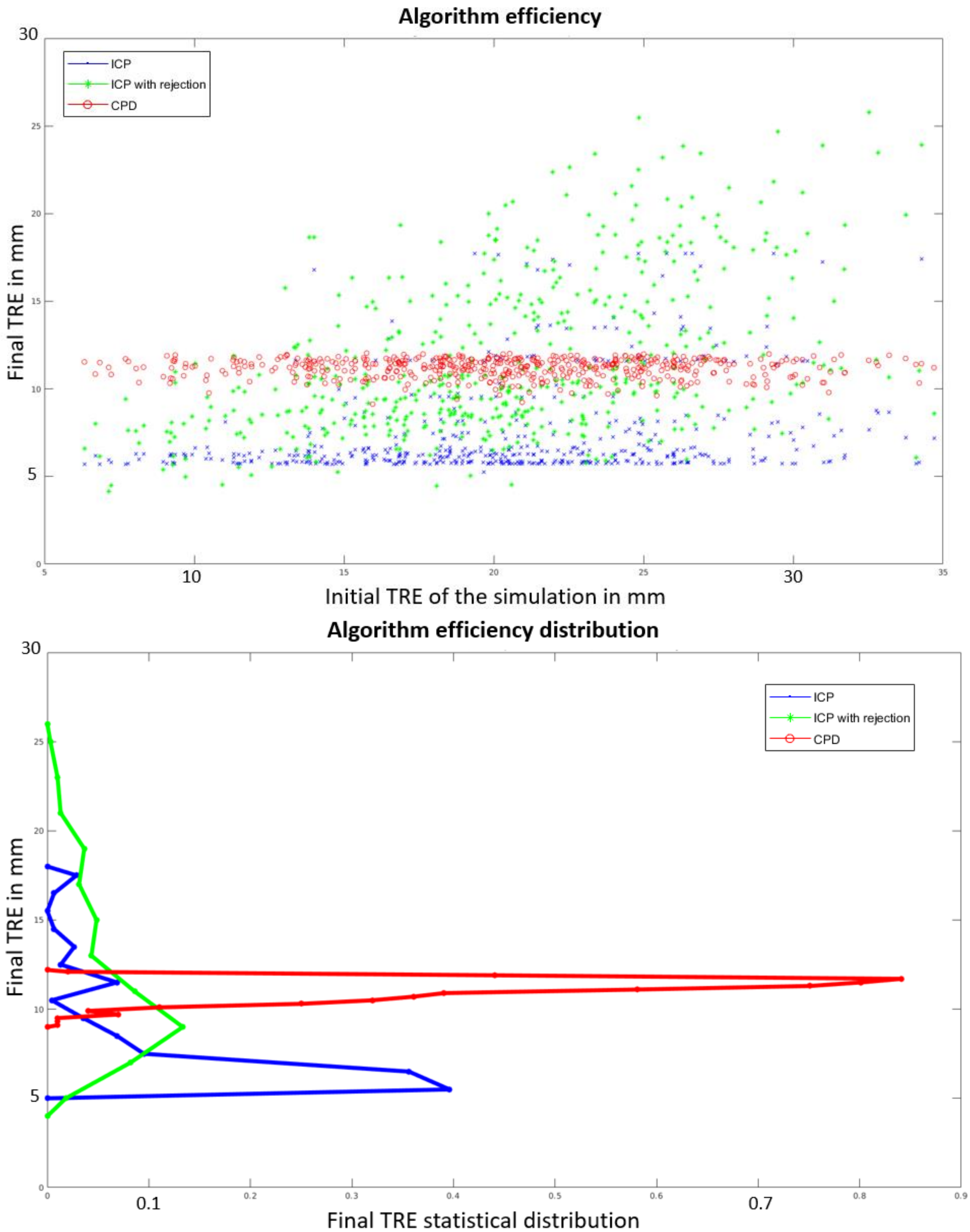


Figure 4.27 : From top to bottom, values of the final TRE for 500 tests, on L1 porcine datasets of 2000 points, with 40 iterations given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green corresponds to robust ICP and red to CPD.

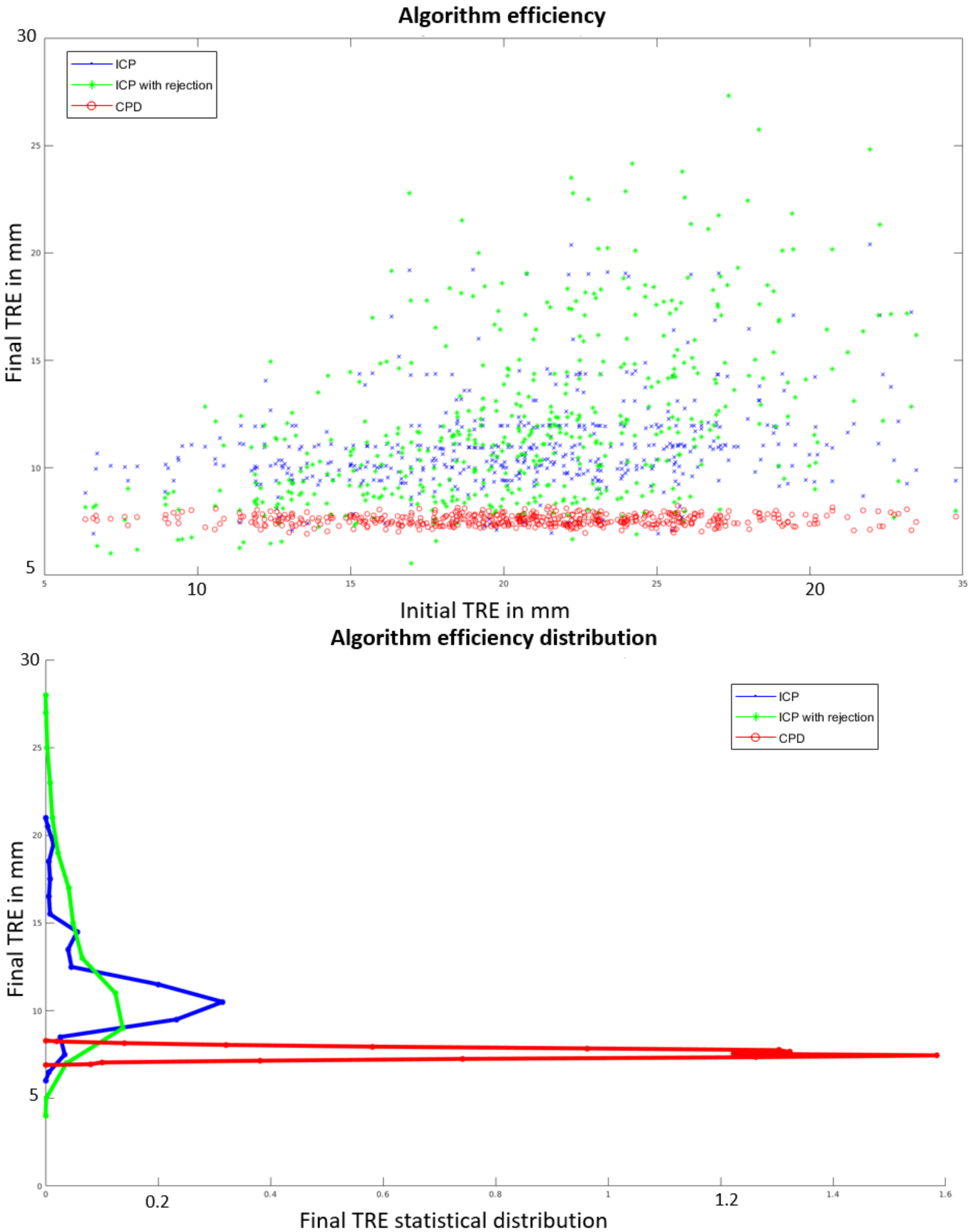


Figure 4.28: For top to bottom, values of the final TRE for 500 tests, on L2 porcine datasets of 2000 points, with 40 iterations given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green correspond to classic ICP, green corresponds to robust ICP and red to CPD.

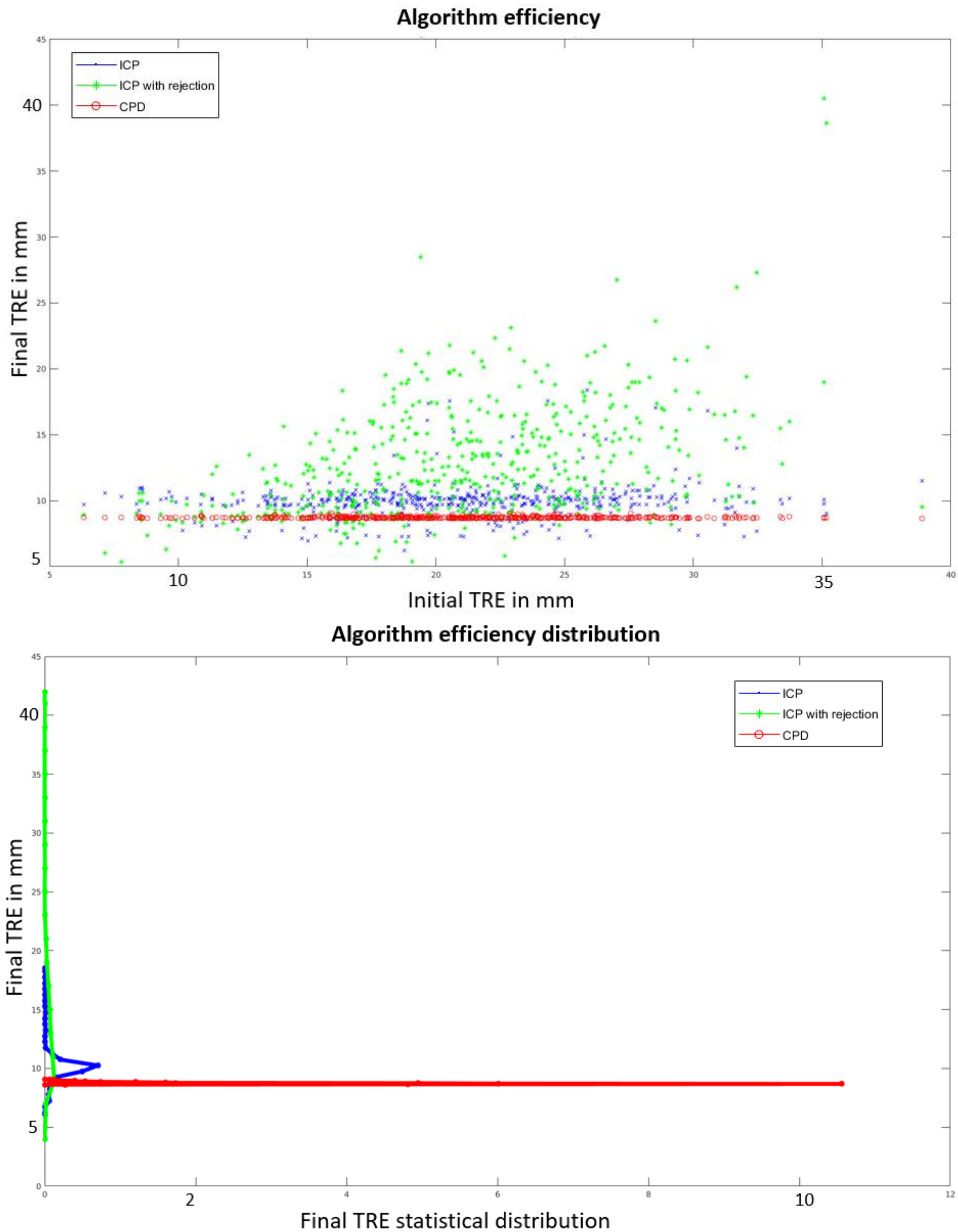


Figure 4.29: From left to right, values of the final TRE for 500 tests, on L3 porcine datasets of 2000 points, with 40 iterations given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green corresponds to robust ICP and red to CPD.

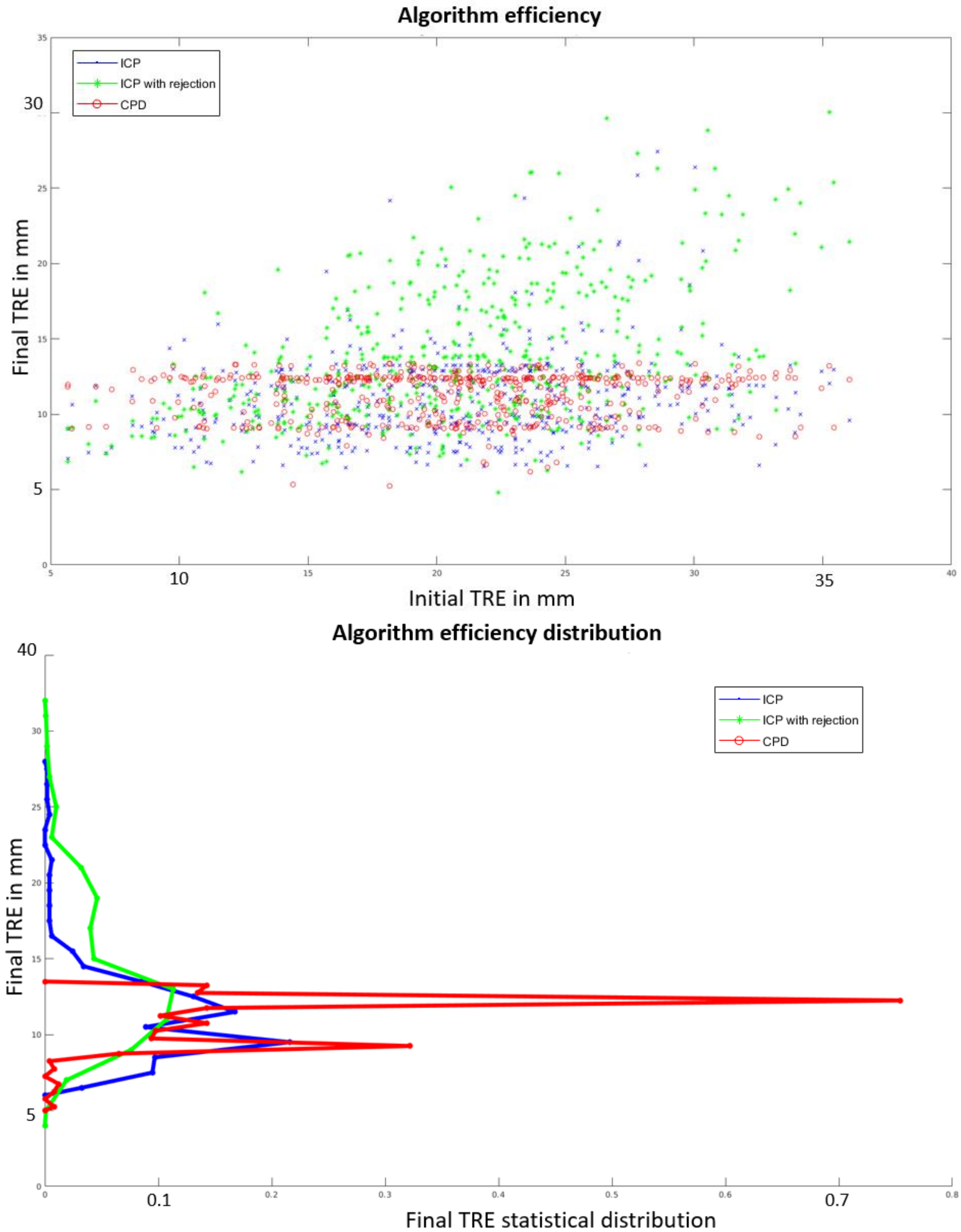


Figure 4.30 From left to right, values of the final TRE for 500 tests, on L4 porcine datasets of 2000 points, with 40 iterations given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green corresponds to robust ICP and red to CPD.

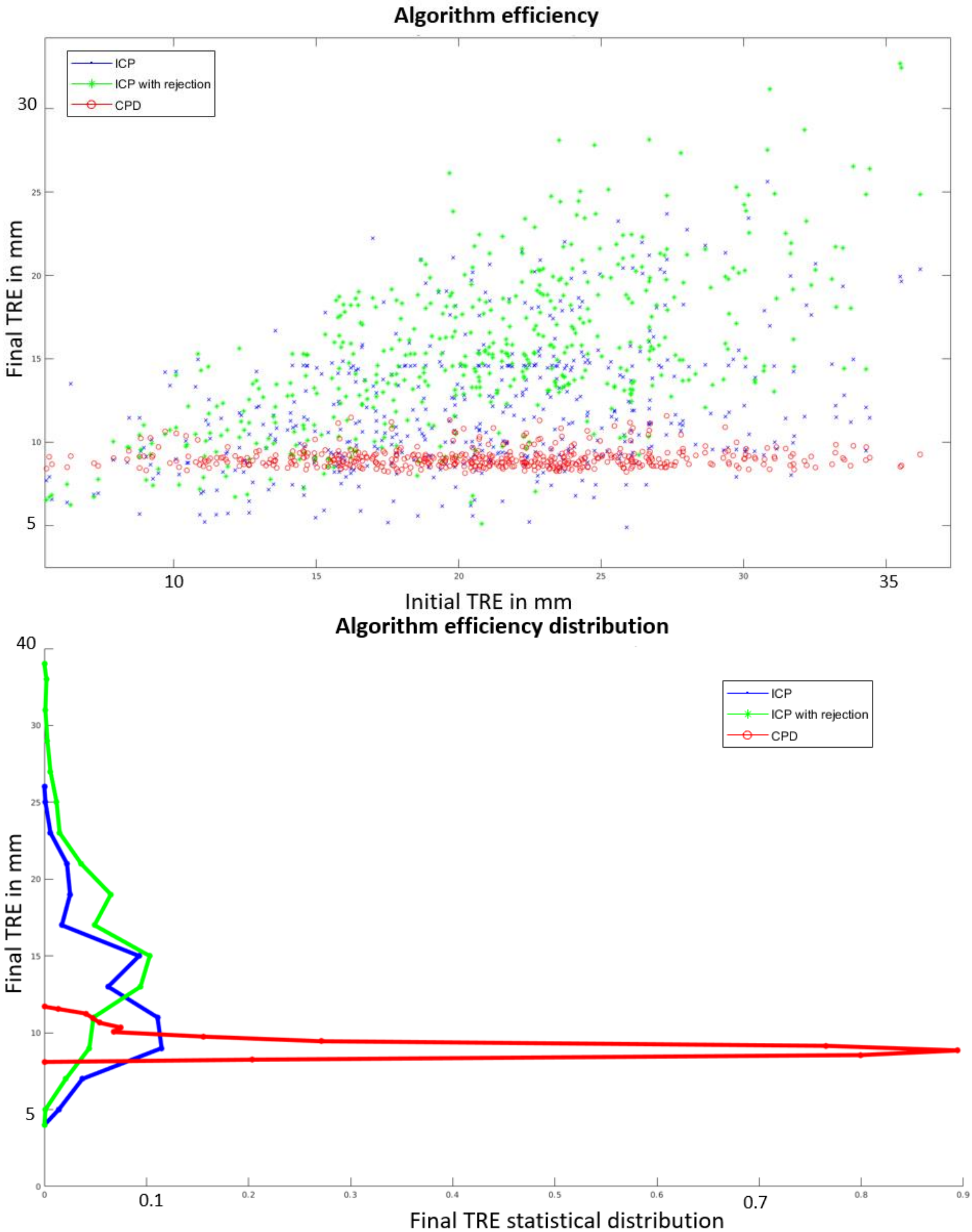


Figure 4.31 : From left to right, values of the final TRE for 500 tests, on L5 porcine datasets of 2000 points, with 40 iterations given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green corresponds to robust ICP and red to CPD

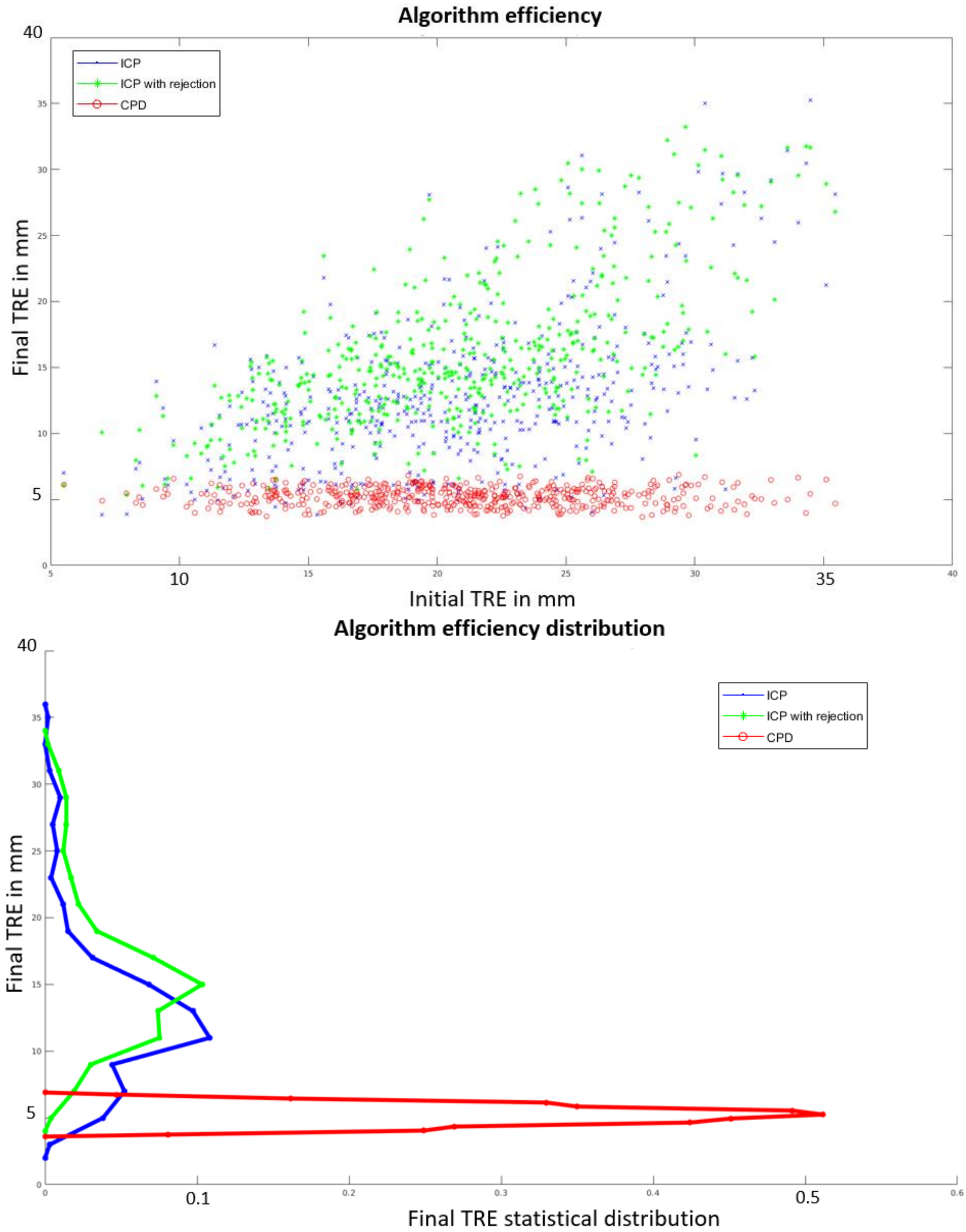


Figure 4.32: From left to right, values of the final TRE for 500 tests, on L6 porcine datasets of 2000 points, with 40 iterations and 20% rejection rate, given the initial TRE, and its distribution for the three algorithms. Blue points and curve correspond to classic ICP, green corresponds to robust ICP and red to CPD

.

About the runtime of each algorithm, ICP and robust ICP have approximately the same duration of a few tenths of a second per registration, while CPD has a duration of a few seconds per registration. Both durations are perfectly tolerable in the context of clinical use.

Chapter 5. Discussions

5.1. About ellipsoids

There were two goals to simulating points on an ellipsoid.

The first one, code validation, showed that the three algorithms were indeed implemented correctly in MATLAB, allowing for the computation of as many points cloud registrations as wished. Figure 4.2 shows that, at least visually, the registration operates correctly.

The second goal was the verification that the two cost functions, RMSE and NLL, had a non-increasing behavior. Figure 4.3 confirms it.

5.2. About the Sawbones lumbar phantom

The first objective of registering two subsampling of the same CT scan segmentation of the lumbar phantom is to verify, incrementally, that the registration algorithms converge for more complex shapes with more noise.

The evolution of RMSE as a function of the transformation parameters show that the cost function is quite smooth, with a minimum corresponding to no transformation, as expected. The cost function does not equal zero at the minimum for two reasons. First, two different subsampling were registered together, making exact superposition impossible. Second, the 2 mm standard deviation Gaussian noise added to the points contributes to the RMSE.

The value of the minimal error is, however, closer to 1 mm, which is smaller than the Gaussian noise added because the computation of the cost function includes taking the average of squared distances for all points. This averaging significantly decreases the noise.

The efficiency of the registration algorithms applied to the Sawbones phantom is measured in part in the variation of the post-registration TRE for multiple transformation values. One could expect these curves to be U shaped as well. This is not what Figure 4.5 shows. The range of transformation that covers clinical needs is small enough for the algorithms not to fail, keeping a final TRE oscillating around 1mm, which is coherent with the minimum RMSE value. This shows the robustness of the algorithms regarding each transformation parameter, as well as the precision level that the algorithms can reach. This interpretation was confirmed when the range was extended to 90° rotations and 90 mm translation distance in Figure 4.6, for different point clouds sizes and number of iterations. Our clinical application does not require managing rotation larger than 70°; however, modified ICP methods were explicitly developed to handle these cases robustly. [53]

An assumption was made that the x-axis was representative of the two other axes and that the convergence basin of the three algorithms was approximately the same size for every direction. Figure 4.5 proves it is at least valid for the Sawbones phantom inside the range of interest.

The method used to determine the sufficient number of points in the dataset was more precise than for the sufficient number of iterations. Indeed, the decrease in TRE with the number of points is subject to some variability because the subsamples sizes represent between 0.35% and 7% of the total dataset. Each point of the figure is the average of 5 trials in order to limit the influence of drastic inhomogeneities in the subsampling, which would make the curve artificially spike without burying the information about the noise.

It appears indeed that the average TRE is not monotonous with the number of points. This is simply due to the variability in initial conditions. Each point being an average of 5 trials, if one of the 5 trials shows a strong local minimum of the optimization algorithm, it would be enough to nudge the average error up. The trade-off was done between the smoothing effect of making multiple simulations of the same transformation and the “double-dipping”, the fact that some point are certainly common to multiple simulations and that if the proportion becomes too large, it could bias the results. This will also appear with the porcine data.

Statistical analysis of the mean and standard deviation of each curve in Figure 4.6 and Figure 4.7, as presented in Table 4.1 and Table 4.2, allow for the use of sophisticated statistical tools; however, the main interest of these simulations was to find a suitable value for the parameters N^* and i^*.

The fact that such small numbers of points still achieve submillimeter precision in the registration allows for fast computation.

The final test showing the result of 500 registrations of the same subsampling using the three different methods shows how close their efficiencies are. Their final TRE distributions have a maximum of around 1 mm showing the same level of accuracy, and the width is nearly identical for the three of them showing the same level of precision.

About the computing time, when run on a MacBook Pro mid2012 (Apple, Cupertino, CA) with 2.5GHz 2 cores processor, 4Gb RAM and MatLab 2019a (Mathworks, Natick, MA), even though one iteration of CPD takes ~100ms while one iteration of ICP takes ~10ms, CPD converges much faster than ICP, for a total time of 5-10s for all three methods. CPD tends to be closer to 10 seconds while ICP is closer to 5 seconds, but this is not a difference of large enough magnitude to discriminate between the two.

5.3. About porcine cadavers

The first difference between the Sawbones phantom and the porcine experiment is that the shape of the US segmentation volume is not similar to the CT segmentation volume, whereas, during experiment 2, the two shapes coregistered came from the same dataset. The other fundamental difference is that the noise of each dataset is not isotropic and Gaussian anymore. It is mainly directed perpendicular to the bone surface, and in the case of ultrasound, since the volume is a reconstruction of a stack of 2D scans, the noise distribution inside each 2D image is not the same as the noise distribution across the stack of US slices. When comparing Figure 4.7, which was made with Sawbones data with Figure 4.10, which is equivalent but based on L3 CT and US data, the baseline of TRE is not around 1 mm but slightly below 10 mm, which is substantially larger. The second difference is that the basin of convergence of the algorithms reduced drastically to ±30° for the rotation around the x-axis. It can be noted, however, that the minimum value of the TRE is reached quite consistently with datasets of 1000 points even though it represents 1.2% of the CT dataset and 1.7% of the US dataset, allowing once again fast computation, across all three methods.

Unlike for the Sawbones phantom, where the rejection rate did not seem to make a significant difference in the efficiency of robust ICP, Figure 4.11 illustrates very clearly the trade-off described in section 3.2.

As mentioned in the discussions surrounding Figure 4.15to Figure 4.20 it would appear that the RMSE, which is the cost function optimized by the classic ICP and the robust ICP algorithms, is not at a global minimum when applying the gold standard transformation. This would explain the horizontal offsets of the curve minima, especially visible for the transformation parameters.

This would mean that one of the following steps is introducing bias in the registration: acquisition of the US scans, acquisition of CT scan, infrared tracking of the US probe, segmentation of US scan, segmentation of CT scan, registration of the US data in the patient frame or computation of the gold standard.

The US acquisition procedure is precisely described in Dr. Yan’s thesis, and the CT acquisition was made following the clinical standards of the Montreal Neurological Institute and Hospital.

In order to test the tracking information, the registration of US data in the patient frame, and the application of the gold standard to CT data, a visual comparison of the application of the gold standard landmark-based registration to the original scans was made. The superposition of both scan information registered using the gold transform can be seen in Figure 5.1. The two scans appear to be correctly registered together.

The manual segmentation of data was repeated multiple times using different approaches but to no avail. It is a viable possibility, however, that the segmentations suffered from some flaw. This is quite difficult to verify, and visual evaluation of the resulting segmented volumes was satisfactory. Nonetheless, it highlights the central role of segmentation in the global procedure. In order for the total procedure to be less than a few minutes long, as requested by the neurosurgeon community [4], automatic segmentation will be unavoidable, and the present situation shows the paramount importance of its accuracy.

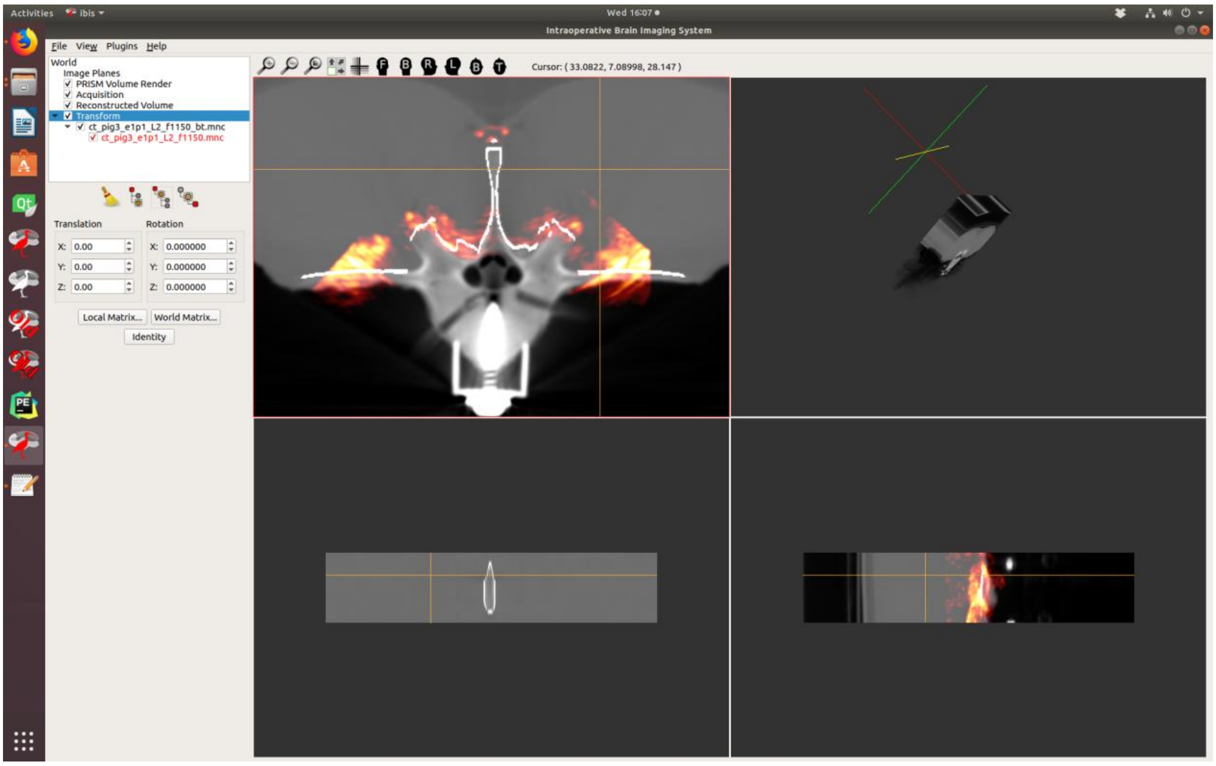


Figure 5.1: comparison of the US and CT scans registered using the gold standard of registration

The highly irregular shapes that take the RMSE for some transformation parameters, such as the rotational parameters for the scans of L4 and L5, shows that this bias in the segmented scans may make the algorithms diverge, or being trapped in a local minimum of the cost function.

The minimum RMSE value of the translation parameters is around 4 or 4.5 mm. When computing the square root of the sum of those distances squared, the total distance error that the RMSE converges to lies between 7.5mm and 10mm, which is consistent with the minimum values present in the TRE graphs.

Even though the accuracy has significantly decreased between Sawbones phantom and porcine cadavers, the robustness and precision of the algorithms seem at least partially preserved. Figure 4.21to Figure 4.26 shows that in most cases, the TRE remains inside the basin of convergence of the algorithms. These zones of convergence vary from one vertebra to the other; however, except for vertebra L4, the curves seem very regular, and the central flat part is analogous to what was observed with the Sawbones phantom. As can be seen in Figure 4.21 to Figure 4.26, the TRE for the CPD algorithm rarely increases above 15mm and stays closer to a constant 10 mm for most tests.

Unlike during Experiment 2, where the three algorithms achieved nearly identical results, on the challenging datasets that are the porcine scans, CPD demonstrates its precision and robustness. Specifically, Figure 4.27 in to Figure 4.30, the thinness of the distribution demonstrates strong robustness regarding the initial positioning. Indeed, the vertical width of the TRE distribution in Figure 4.11 is of the order of 1mm. In comparison, even though the third dataset is much noisier and globally less accurate, the fluctuations of the CPD algorithm error stays in this millimetric interval for L1, L4, L5 and L6, submillimetric in the case of L2 and L3. One discrepancy can be noted, however, between Figure 4.28 and Figure 4.32. The value to which CPD converges appears to be around 10 mm in Figure 4.30 much above ICP and robust ICP, whereas Figure 4.29 shows that CPD converges to lower values. This could be an effect of a local minimum of the NLL function, as it is, for example, with L4 in Figure 4.30. Looking closely at the distributions of ICP and robust ICP also shows the appearance of plateaus, but they are much less visible due to the importance of the noise and bias in the datasets, but they can be observed for L2, L4 and L5, in Figure 4.28, Figure 4.30, and Figure 4.31

5.4. Future works

One line of investigation which was already suggested in this work is the development of a consensus regarding robust and fast segmentation algorithms for US and CT data. In order to have a complete system that satisfies clinical requirements, each step of the procedure must be optimized independently. Substantial literature was written on rigid registration, but other steps of the protocol have not been as thoroughly examined.

As the literature review illustrates, an excessively large number of algorithms for rigid registration were developed in the last 20 to 30 years. Some of them were tested on real images, others on synthetic datasets, others, but more rarely, on medical images. A tremendously significant advance for the domain would be to implement every method in the same computing framework, and systematically test them on real datasets, clinical application per clinical application. Once these comparisons are rigorously carried, standards can be agreed on that will facilitate the propagation of the method and transmission to the clinicians the least in contact with research facilities.

Non-rigid registration has also seen substantial developments in the last few years, and specifically, the emergence of Machine Learning has redefined expectations about the efficiency of image analysis. Trying to apply some form of Machine Learning to the problem of real-time multimodal registration could lead to a new paradigm in the domain.

Conclusion

The goal of this thesis was to compare three promising point-based registration algorithms for US-CT registration of vertebrae during spinal fusion surgery in increasingly realistic contexts. The Coherent Point Drift algorithm demonstrated precision, strong robustness regarding noise, and initial misalignment. Further validations on better quality data, human cadavers, and live patients will be necessary in order to incorporate it into a fully automatic process. This research was carried in the hope of providing the surgical community with a cheap, fast, and radiation-free registration technique, which would be implemented inside the broader scope of an Image-Guided System for neurosurgery.

Bibliography

[1] W.C. Johnson, A. Seifi, Trends of the neurosurgical economy in the United States, J. Clin. Neurosci. 53 (2018) 20–26. doi:10.1016/j.jocn.2018.04.041.

[2] S.S. Rajaee, H.W. Bae, L.E.A. Kanim, R.B. Delamarter, Spinal fusion in the United States: Analysis of trends from 1998 to 2008, Spine (Phila. Pa. 1976). 37 (2012) 67–76. doi:10.1097/BRS.0b013e31820cccfb.

[3] R.A. Deyo, A. Nachemson, S.K. Mirza, Spinal-Fusion Surgery—The Case for Restraint, Spine J. 4 (2004) S138–S142. doi:10.1016/j.spinee.2004.08.001.

[4] K. Cleary, J. Anderson, M. Brazaitis, G. Devey, A. DiGioia, M. Freedman, D. Grönemeyer, C. Lathan, H. Lemke, D. Long, S.K. Mun, R. Taylor, Final report of the technical requirements for image-guided spine procedures workshop, Comput. Aided Surg. 5 (2000) 180–215. doi:10.1002/1097-0150(2000)5:3<180::AID-IGS6>3.0.CO;2-C.

[5] M. Beck, T. Mittlmeier, P. Gierer, C. Harms, G. Gradl, Benefit and accuracy of intraoperative 3D-imaging after pedicle screw placement: A prospective study in stabilizing thoracolumbar fractures, Eur. Spine J. 18 (2009) 1469–1477. doi:10.1007/s00586-009-1050-5.

[6] S. Knafo, E. Mireau, S. Bennis, B. Baussart, S. Aldea, S. Gaillard, Operative and perioperative durations in O-arm vs C-arm fluoroscopy for lumbar instrumentation, Oper. Neurosurg. 14 (2018) 273–278. doi:10.1093/ons/opx142.

[7] J.L. Prince, J.M. Links, Medical Imaging, Signals and Systems, 2nd Ed, Prentice Hall, 2015.

[8] P.. Besl, N.D. McKay, A method for registration of 3-D shapes, IEEE Trans. Pattern Anal. Mach. Intell. 14 (1992) 239–256. doi:10.1109/34.121791.

[9] A. Myronenko, X. Song, Point Set Registration : Coherent Point Drift, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010) 2262–2275. doi:10.1109/TPAMI.2010.46.

[10] C.X.B. Yan, B. Goulet, D. Tampieri, D.L. Collins, Ultrasound-CT registration of vertebrae without reconstruction, Int. J. Comput. Assist. Radiol. Surg. 7 (2012) 901–909. doi:10.1007/s11548-012-0771-9.

[11] T. Barclay, Lumbar Vertebrae - Anatomy Pictures and Information, Innerbody. (2019). https://www.innerbody.com/anatomy/skeletal/lumbar-vertebrae-lateral (accessed August 15, 2019).

[12] C. Yan, Ultrasound-CT Registration of Vertebrae for Image-Guided Spinal Fusion Surgeries, 2013. http://digitool.library.mcgill.ca/webclient/StreamGate?folder_id=0&dvs=1542151834533~649.

[13] T. Tjardes, S. Shafizadeh, D. Rixen, T. Paffrath, B. Bouillon, E.S. Steinhausen, H. Baethis, Image-guided spine surgery: State of the art and future directions, Eur. Spine J. 19 (2010) 25–45. doi:10.1007/s00586-009-1091-9.

[14] R.C. Sasso, N.M. Best, E.A. Potts, Percutaneous computer-assisted translaminar facet screw: an initial human cadaveric study, Spine J. 5 (2005) 515–519. doi:10.1016/j.spinee.2005.03.016.

[15] P.A. Helm, R. Teichman, S.L. Hartmann, D. Simon, Spinal Navigation and Imaging: History, Trends, and Future, IEEE Trans. Med. Imaging. 34 (2015) 1738–1746. doi:10.1109/TMI.2015.2391200.

[16] C.X.B. Yan, B. Goulet, J. Pelletier, S.J.S. Chen, D. Tampieri, D.L. Collins, Towards accurate, robust and practical ultrasound-CT registration of vertebrae for image-guided spine surgery, Int. J. Comput. Assist. Radiol. Surg. 6 (2011) 523–537. doi:10.1007/s11548-010-0536-2.

[17] S. Pieper, M. Halle, R. Kikinis, 3D Slicer, (2005) 632–635. doi:10.1109/isbi.2004.1398617.

[18] A. Enquobahrie, P. Cheng, K. Gary, L. Ibanez, D. Gobbi, F. Lindseth, Z. Yaniv, S. Aylward, J. Jomier, K. Cleary, The Image-Guided Surgery Toolkit IGSTK: An open source C++ software toolkit, J. Digit. Imaging. 20 (2007) 21–33. doi:10.1007/s10278-007-9054-3.

[19] E. Van De Kelft, F. Costa, D. Van Der Planken, F. Schils, A prospective multicenter registry on the accuracy of pedicle screw placement in the thoracic, lumbar, and sacral levels with the use of the O-arm imaging system and stealthstation navigation, Spine (Phila. Pa. 1976). 37 (2012). doi:10.1097/BRS.0b013e318271b1fa.

[20] S. Drouin, A. Kochanowska, M. Kersten-Oertel, I.J. Gerard, R. Zelmann, D. De Nigris, S. Bériault, T. Arbel, D. Sirhan, A.F. Sadikot, J.A. Hall, D.S. Sinclair, K. Petrecca, R.F. DelMaestro, D.L. Collins, IBIS: an OR ready open-source platform for image-guided neurosurgery, Int. J. Comput. Assist. Radiol. Surg. 12 (2017) 363–378. doi:10.1007/s11548-016-1478-0.

[21] C. Askeland, O.V. Solberg, J.B.L. Bakeng, I. Reinertsen, G.A. Tangen, E.F. Hofstad, D.H. Iversen, C. Våpenstad, T. Selbekk, T. Langø, T.A.N. Hernes, H. Olav Leira, G. Unsgård, F. Lindseth, CustusX: an open-source research platform for image-guided therapy, Int. J. Comput. Assist. Radiol. Surg. 11 (2016) 505–519. doi:10.1007/s11548-015-1292-0.

[22] L. Mercier, R.F. Del Maestro, K. Petrecca, A. Kochanowska, S. Drouin, C.X.B. Yan, A.L. Janke, S.J.S. Chen, D.L. Collins, New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: System description and validation, Int. J. Comput. Assist. Radiol. Surg. 6 (2011) 507–522. doi:10.1007/s11548-010-0535-3.

[23] I.J. Gerard, J.A. Hall, K. Mok, D.L. Collins, New protocol for skin landmark registration in image-guided neurosurgery: Technical note, Clin. Neurosurg. 11 (2015) 376–381. doi:10.1227/NEU.0000000000000868.

[24] T. Ault, M.W. Siegel, Frameless patient registration using ultrasonic imaging: A preliminary study, Comput. Aided Surg. 1 (1995) 94–102. doi:10.3109/10929089509106320.

[25] I. Hacihaliloglu, R. Abugharbieh, A.J. Hodgson, R.N. Rohling, Bone Surface Localization in Ultrasound Using Image Phase-Based Features, Ultrasound Med. Biol. 35 (2009) 1475–1487. doi:10.1016/j.ultrasmedbio.2009.04.015.

[26] I. Hacihaliloglu, Enhancement of bone shadow region using local phase-based ultrasound transmission maps, Int. J. Comput. Assist. Radiol. Surg. (2017) 1–10. doi:10.1007/s11548-017-1556-y.

[27] I. Hacihaliloglu, A. Rasoulian, R.N. Rohling, P. Abolmaesumi, Local phase tensor features for 3-D ultrasound to statistical shape+pose spine model registration, IEEE Trans. Med. Imaging. 33 (2014) 2167–2179. doi:10.1109/TMI.2014.2332571.

[28] B. Brendel, S. Winter, A. Rick, M. Stockheim, H. Ermert, Registration of 3D CT and ultrasound datasets of the spine using bone structures, Comput. Aided Surg. 7 (2002) 146–155. doi:10.1002/igs.10038.

[29] F. Alam, S.U. Rahman, Intrinsic registration techniques for medical images: A state-of-the-art review, J. Postgrad. Med. Inst. 30 (2016) 119–132.

[30] Y. Kang, K. Engelke, W.A. Kalender, A New Accurate and Precise 3-D Segmentation.pdf, IEEE Trans. Med. Imaging. 22 (2003) 586–598.

[31] F.P.M. Oliveira, J.M.R.S. Tavares, Medical image registration: a review, Comput.

Methods Biomech. Biomed. Engin. 17 (2014) 73–93. doi:10.1080/10255842.2012.670855.

[32] L.G. Brown, A survey of image registration techniques, ACM Comput. Surv. 24 (1992) 325–376. doi:10.1145/146370.146374.

[33] J.B. Maintz, M.A. Viergever, A survey of medical image registration., Med. Image Anal. 2 (1998) 1–36. http://www.ncbi.nlm.nih.gov/pubmed/10638851.

[34] M.A. Viergever, J.B.A. Maintz, S. Klein, K. Murphy, M. Staring, J.P.W. Pluim, A survey of medical image registration – under review, Med. Image Anal. 33 (2016) 140–144. doi:10.1016/j.media.2016.06.030.

[35] D.L.G. Hill, P.G. Batchelor, M. Holden, D.J. Hawkes, Medical image registration, Phys. Med. Biol. 46 (2001) R1–R45. doi:10.1088/0031-9155/46/3/201.

[36] V.R.S. Mani, S. Arivazhagan, Survey of Medical Image Registration, J. Biomed. Eng. Technol. 1 (2013) 8–25. doi:10.12691/jbet-1-2-1.

[37] S. Schumann, Computational Radiology for Orthopaedic Interventions, 23 (2016). doi:10.1007/978-3-319-23482-3.

[38] F. Langlotz, L.P. Nolte, Technical Approaches to Computer-Assisted Orthopedic Surgery, Eur. J. Trauma. 30 (2004) 1–11. doi:10.1007/s00068-004-1374-0.

[39] E. Ferrante, N. Paragios, Slice-to-volume medical image registration: A survey, Med. Image Anal. 39 (2017) 101–123. doi:10.1016/j.media.2017.04.010.

[40] J. Alison Noble, N. Navab, H. Becher, Ultrasonic image analysis and imageguided interventions, Interface Focus. 1 (2011) 673–685. doi:10.1098/rsfs.2011.0025.

[41] L.-P. Amiot, H. Labelle, J.A. DeGuise, M. Sati, P. Brodeur, C.-H. Rivard, Computer-Assisted Pedicle Screw Fixation- A Feasibility Study, Spine (Phila. Pa. 1976). 20 (1995) 1208–1212. doi:10.1097/00007632-199505150-00019.

[42] A. Brounstein, I. Hacihaliloglu, P. Guy, A. Hodgson, R. Abugharbieh, Towards real-time 3D US to CT bone image registration using phase and curvature feature based GMM matching, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 6891 LNCS (2011) 235–242. doi:10.1007/978-3-642-23623-5_30.

[43] C.R. Meyer, J.L. Boes, B. Kim, P.H. Bland, G.L. Lecarpentier, J.B. Fowlkes, M.A. Roubidoux, P.L. Carson, Semiautomatic registration of volumetric ultrasound scans., Ultrasound Med. Biol. 25 (1999) 339–47.

http://www.ncbi.nlm.nih.gov/pubmed/10374978.

[44] F. Pomerleau, F. Colas, R. Siegwart, A Review of Point Cloud Registration Algorithms for Mobile Robotics, Found. Trends Robot. 4 (2015) 1–104. doi:10.1561/2300000035.

[45] Y. Chen, G. Medioni, Object modelling by registration of multiple range images, Image Vis. Comput. 10 (1992) 145–155. doi:10.1016/0262-8856(92)90066-C.

[46] G. Champleboux, S. Lavallee, R. Szeliski, L. Brunie, From accurate range imaging sensor calibration to accurate model-based 3D object localization, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 1992-June (1992) 83–89. doi:10.1109/CVPR.1992.223223.

[47] A.W. Fitzgibbon, Robust registration of 2D and 3D point sets, Image Vis. Comput. 21 (2003) 1145–1153. doi:10.1016/j.imavis.2003.09.004.

[48] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, Int. J. Comput. Vis. 13 (1994) 119–152. doi:10.1007/BF01427149.

[49] S. Rusinkiewicz, M. Levoy, Efficient variants of the ICP algorithm, in: Proc. Third Int. Conf. 3-D Digit. Imaging Model., IEEE Comput. Soc, 2001: pp. 145–152. doi:10.1109/IM.2001.924423.

[50] B. Ma, R.E. Ellis, Robust registration for computer-integrated orthopedic surgery: Laboratory validation and clinical experience, Med. Image Anal. 7 (2003) 237–250. doi:10.1016/S1361-8415(02)00133-0.

[51] F. Pomerleau, F. Colas, R. Siegwart, S. Magnenat, Comparing ICP variants on real-world data sets: Open-source library and experimental protocol, Auton. Robots. 34 (2013) 133–148. doi:10.1007/s10514-013-9327-2.

[52] M. Magnusson, The Three-Dimensional Normal-Distributions Transform — an Efficient Representation for Registration, Surface Analysis, and Loop Detection, Örebro University, 2013. http://www.aass.oru.se/Research/Learning/publications/2009/Magnusson_2009-Doctoral_Thesis-3D_NDT.pdf.

[53] T. Petricek, T. Svoboda, Point cloud registration from local feature correspondences—Evaluation on challenging datasets, PLoS One. 12 (2017) e0187943. doi:10.1371/journal.pone.0187943.

[54] H. Chui, A. Rangarajan, A feature registration framework using mixture models, in: Proc. IEEE Work. Math. Methods Biomed. Image Anal. MMBIA-2000 (Cat.

No.PR00737), IEEE Comput. Soc, n.d.: pp. 190–197. doi:10.1109/MMBIA.2000.852377.
[55] A. Segal, D. Haehnel, S. Thrun, Generalized-ICP, in: Robot. Sci. Syst. V, Robotics: Science and Systems Foundation, 2009. doi:10.15607/RSS.2009.V.021.
[56] F. Rodriguez y Baena, T. Hawke, M. Jakopec, A bounded iterative closest point method for minimally invasive registration of the femur, Proc. Inst. Mech. Eng. Part H J. Eng. Med. 227 (2013) 1135–1144. doi:10.1177/0954411913500948.
[57] Weiguo Xie, L. Nolte, Guoyan Zheng, ECM versus ICP for point registration, in: 2011 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE, 2011: pp. 2131–2135. doi:10.1109/IEMBS.2011.6090398.
[58] R. Horaud, F. Forbes, M. Yguel, G. Dewaele, J. Zhang, Rigid and Articulated Point Registration with Expectation Conditional Maximization, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 587–602. doi:10.1109/TPAMI.2010.94.
[59] Y. He, B. Liang, J. Yang, S. Li, J. He, An Iterative Closest Points Algorithm for Registration of 3D Laser Scanner Point Clouds with Geometric Features, Sensors. 17 (2017) 1862. doi:10.3390/s17081862.
[60] S. Du, Y. Xu, T. Wan, H. Hu, S. Zhang, G. Xu, X. Zhang, Robust iterative closest point algorithm based on global reference point for rotation invariant registration, PLoS One. 12 (2017) e0188039. doi:10.1371/journal.pone.0188039.
[61] G.C. Sharp, S. Member, S.W. Lee, D.K. Wehe, ICP Registration Using Invariant Features, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 90–102. doi:10.1109/34.982886.
[62] J. Yang, S. Member, H. Li, Go-ICP : A Globally Optimal Solution to 3D ICP Point-Set Registration, IEEE Trans. Pattern Anal. Mach. Intell. 38 (2016) 2241–2254. doi:10.1109/TPAMI.2015.2513405.
[63] P. Wang, P. Wang, Z.G. Qu, Y.H. Gao, Z.K. Shen, A refined coherent point drift (CPD) algorithm for point set registration, Sci. China Inf. Sci. 54 (2011) 2639–2646. doi:10.1007/s11432-011-4465-7.
[64] L. Maier-Hein, A.M. Franz, T.R. dos Santos, M. Schmidt, M. Fangerau, H. Meinzer, J.M. Fitzpatrick, Convergent Iterative Closest-Point Algorithm to Accomodate Anisotropic and Inhomogenous Localization Error, IEEE Trans. Pattern Anal. Mach. Intell. 34 (2012) 1520–1532. doi:10.1109/TPAMI.2011.248.
[65] H. Mora, J.M. Mora-Pascual, A. García-García, P. Martínez-González, Computational

Analysis of Distance Operators for the Iterative Closest Point Algorithm, PLoS One. 11 (2016) e0164694. doi:10.1371/journal.pone.0164694.

[66] S.D. Billings, E.M. Boctor, R.H. Taylor, Iterative Most-Likely Point Registration (IMLP): A Robust Algorithm for Computing Optimal Shape Alignment, PLoS One. 10 (2015) e0117688. doi:10.1371/journal.pone.0117688.

[67] Z. Cao, S. Pan, R. Li, R. Balachandran, J.M. Fitzpatrick, W.C. Chapman, B.M. Dawant, Registration of medical images using an interpolated closest point transform: method and validation, Med. Image Anal. 8 (2004) 421–427. doi:10.1016/j.media.2004.01.002.

[68] N.J. Mitra, N. Gelfand, H. Pottmann, L. Guibas, Registration of point cloud data from a geometric optimization perspective, in: Proc. 2004 Eurographics/ACM SIGGRAPH Symp. Geom. Process. - SGP ’04, ACM Press, New York, New York, USA, 2004: p. 22. doi:10.1145/1057432.1057435.

[69] Bing Jian, B.C. Vemuri, Robust Point Set Registration Using Gaussian Mixture Models, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 1633–1645. doi:10.1109/TPAMI.2010.223.

[70] C.V. Stewart, Chia-Ling Tsai, B. Roysam, The dual-bootstrap iterative closest point algorithm with application to retinal image registration, IEEE Trans. Med. Imaging. 22 (2003) 1379–1394. doi:10.1109/TMI.2003.819276.

[71] L. Evans, J.D. Olson, Y. Cai, X. Fan, K.D. Paulsen, D.W. Roberts, S. Ji, S.S. Lollis, Stereovision Co-Registration in Image-Guided Spinal Surgery: Accuracy Assessment Using Explanted Porcine Spines, Oper. Neurosurg. 15 (2018) 686–691. doi:10.1093/ons/opy023.

[72] M.H. Moghari, P. Abolmaesumi, Point-Based Rigid-Body Registration Using an Unscented Kalman Filter, IEEE Trans. Med. Imaging. 26 (2007) 1708–1728. doi:10.1109/TMI.2007.901984.

[73] A.P. King, K.S. Rhode, Y. Ma, C. Yao, C. Jansen, R. Razavi, G.P. Penney, Registering Preprocedure Volumetric Images With Intraprocedure 3-D Ultrasound Using an Ultrasound Imaging Model, IEEE Trans. Med. Imaging. 29 (2010) 924–937. doi:10.1109/TMI.2010.2040189.

[74] M. Pan, J. Tang, Q. Rong, F. Zhang, Medical image registration using modified iterative closest points, Int. j. Numer. Method. Biomed. Eng. 27 (2011) 1150–1166. doi:10.1002/cnm.1421.

[75] N. Gelfand, N.J. Mitra, L. Guibas, H. Pottmann, Robust Global Registration, in: Proc. Third Eurographics Symp. Geom. Process., 2005: p. 197.

[76] J.L. Bentley, Multidimensional binary search trees used for associative searching, Commun. ACM. 18 (1975) 509–517. doi:10.1145/361002.361007.

[77] J.M. Fitzpatrick, J.B. West, The distribution of target registration error in rigid-body point-based registration, IEEE Trans. Med. Imaging. 20 (2001) 917–927. doi:10.1109/42.952729.

[78] H. Kjer, J. Wilm, Evaluation of surface registration algorithms for PET motion correction, Technical University of Denmark, 2010. http://etd.dtu.dk/thesis/262855/.

[79] L. Greengard, J. Strain, The Fast Gauss Transform, SIAM J. Sci. Stat. Comput. 12 (1991) 79–94. doi:10.1137/0912004.

[80] pcregistercpd MATLAB Documentation, MathWorks, Inc. (2018). https://www.mathworks.com/help/vision/ref/pcregistercpd.html.

[81] E.M.A. Anas, A. Seitel, A. Rasoulian, P.S. John, D. Pichora, K. Darras, D. Wilson, V.A. Lessoway, I. Hacihaliloglu, P. Mousavi, R. Rohling, P. Abolmaesumi, Bone enhancement in ultrasound using local spectrum variations for guiding percutaneous scaphoid fracture fixation procedures, Int. J. Comput. Assist. Radiol. Surg. 10 (2015). doi:10.1007/s11548-015-1181-6.

[82] I. Busscher, J.J.W. Ploegmakers, G.J. Verkerke, A.G. Veldhuizen, Comparative anatomical dimensions of the complete human and porcine spine, Eur. Spine J. 19 (2010) 1104–1114. doi:10.1007/s00586-010-1326-9.

[83] Gold standard, Oxford English Dict. (n.d.). https://www.lexico.com/en/definition/gold_standard (accessed August 15, 2019).

[84] Ground Truth, Oxford English Dict. (n.d.). https://www.lexico.com/en/definition/ground_truth.

[85] E. Versi, “Gold standard” is an appropriate term., BMJ. 305 (1992) 187–187. doi:10.1136/bmj.305.6846.187-b.

[86] R.M. Comeau, A.F. Sadikot, A. Fenster, T.M. Peters, Intraoperative ultrasound for guidance and tissue shift correction in image-guided neurosurgery, Med. Phys. 27 (2000) 787–800. doi:10.1118/1.598942.