SPLINE PROJECTION-BASED VOLUME-TO-IMAGE REGISTRATION

THÈSE N^{O} 2901 (2003)

PRÉSENTÉE À LA FACULTÉ SCIENCES ET TECHNIQUES DE L'INGÉNIEUR

Institut d'imagerie et optique appliquée

SECTION DE MICROTECHNIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES TECHNIQUES

PAR

Slavica JONIĆ

M.Sc. in Electrical Engineering, University of Belgrade, Serbie et Monténégro
et de nationalités serbe

acceptée sur proposition du jury:

Prof. M. Unser, directeur de thèse
Dr M. Bierlaire, rapporteur
Prof. J. M. Carazo, rapporteur
Prof. L.-P. Nolte, rapporteur

Lausanne, EPFL
2003

Spline Projection-Based Volume-to-Image Registration

Slavica Jonić

Thèse N° 2901 (2003)

Thèse présentée à la faculté des sciences et techniques de l'ingénieur
pour l'obtention du grade de docteur ès sciences techniques
et acceptée sur proposition du jury:

Prof. Hannes Bleuler, *président*
Prof. Michael Unser, *directeur de thèse*
Dr. Michel Bierlaire, *rapporteur*
Prof. José María Carazo, *rapporteur*
Prof. Lutz-Peter Nolte, *rapporteur*

École Polytechnique Fédérale de Lausanne—2003

Cover design by Annette Unser
Cover printed by Copy-Tunnel
Printing and binding by Repro-EPFL
Typeset with LaTeX

Available at `http://bigwww.epfl.ch/jonic/`

Abstract

This thesis focuses on the rigid-body registration of a three-dimensional model of an object to a set of its two-dimensional projections. The main contribution is the development of two registration algorithms that use a continuous model of the volume based on splines, either in the space domain or in the frequency domain. This allows for a well-defined gradient of the dissimilarity measure, which is a necessary condition for efficient and accurate registration.

The first part of the thesis contains a review of the literature on volume-to-image registration. Then, we discuss data interpolation in the space domain and in the frequency domain.

The basic concepts of our registration strategy are given in the second part of the thesis. We present a novel one-step approach for fast ray casting to simulate space-based volume projections. We also discuss the use of the central-slice theorem to simulate frequency-based volume projections. Then, we consider the question of the registration robustness. To improve the robustness of the space-based approach, we apply a multiresolution optimization strategy where spline-based data pyramids are processed in coarse-to-fine fashion, which improves speed as well. To improve the robustness of the frequency-based registration, we apply a coarse-to-fine strategy that involves weights in the frequency domain.

In the third part, we apply our space-based algorithm to computer-assisted orthopedic surgery while adapting it to the perspective projection model. We show that the registration accuracy achieved using the orthopedic data is consistent with the current standards. Then, we apply our frequency-based registration to three-dimensional electron-microscopy application. We show that our algorithm can be used to obtain a refined solution with respect to currently available algorithms. The novelty of our approach is in dealing with a continuous space of geometric parameters, contrary to the standard methods which deal with quantized parameters. We conclude that our continuous parameter space leads to better registration accuracy.

Last, we compare the performance of the frequency-based algorithm with that of the space-based algorithm in the context of electron microscopy. With these data, we observe that frequency-based registration algorithm outperforms the space-based one, which we attribute to the suitability of interpolation in the frequency domain when dealing with strictly space-limited data.

Résumé

Nous nous intéressons à la mise en correspondance rigide d'un modèle troidimensionnel avec l'ensemble de ses projections bidimensionnelles. La contribution principale est le développement de deux algorithmes de mise en correspondance qui utilisent un modèle continu du volume basé sur des splines, que ce soit dans le domaine spatial ou fréquentiel. Ceci permet de mieux calculer le gradient de la mesure de dissimilarité, condition nécessaire pour une mise en correspondance précise et efficace.

La première partie de cette thèse fait le point sur la littérature de la mise en correspondance de volumes avec des images. Nous parlons ensuite de l'interpolation de données dans le domaine spatial et fréquentiel.

Les concepts de base de notre stratégie de mise en correspondance sont donnés dans la deuxième partie de la thèse. Nous proposons une nouvelle approche rapide en une étape afin de simuler spatialement les projections de rayons à travers le volume. Nous traitons aussi de l'utilisation du théorème de la coupe centrale afin de simuler les projections des volumes dans le domain des fréquences. Ensuite, nous examinons la robustesse de la mise en correspondance. Afin d'améliorer la robustesse de l'approche spatiale, nous appliquons une stratégie d'optimisation en multirésolution où les pyramides de données, basées sur les splines, sont traitées "de l'échelle la plus grossière à la plus fine", ce qui accélère aussi le processus. Pour améliorer la robustesse de l'approche fréquentielle, nous appliquons une stratégie qui implique la mise en place de poids dans le domaine fréquentiel.

Dans la troisième partie, nous appliquons notre algorithme spatial à la chirurgie orthopédique assistée par ordinateur, tout en l'adaptant à la projection perspective. Nous démontrons que la précision de la mise en correspondance obtenue avec les données orthopédiques est conforme aux résultats standards. Ensuite, nous appliquons notre mise en correspondance fréquentielle aux applications en microscopie électronique troidimensionnelle. Nous montrons que

notre algorithme peut être utilisé pour obtenir une meilleure solution que les algorithmes disponibles actuellement. La nouveauté de notre approche est de manipuler les paramètres géomètriques dans un espace continu, contrairement aux méthodes classiques qui manipulent des paramètres quantifiés. Nous concluons que notre espace continu de paramètres apporte une meilleure précision de mise en correspondance.

Enfin, nous comparons la performance de l'algorithme fréquentiel à celle de l'algorithme spatial dans le contexte de la microscopie électronique. Nous y observons que l'algorithme fréquentiel de mise en correspondance surpasse l'algorithme spatial. On peut expliquer ce fait par l'adéquation de l'interpolation dans le domaine fréquentiel à des données qui sont strictement limitées en espace.

To my parents.

Acknowledgements

I would like to express my deep gratitude to my advisors, Michael Unser and Philippe Thévenaz, for their support and encouragement throughout all these four years. The patience and efforts that Philippe has put in reviewing the drafts of the thesis and of the papers should be especially recognized.

I also thank all the other members of the Biomedical Imaging Group (BIG) who have contributed to the work on this thesis. Special thanks to Arrate Muñoz Barrutia, Jan Kybic, Mathews Jacob, Stefan Horbelt, and Carlos Óscar Sánchez Sorzano for all the nice moments we have spent together. I was very lucky to have Carlos Óscar for my officemate during the last year. I have really enjoyed our collaboration.

I should not forget our collaborators who have provided the data used in the thesis. I thank Catherine El-Bez from Laboratoire d'Analyse Ultrastructurale, Université de Lausanne and Guoyan Zheng from MEM Research Center for Orthopaedic Surgery, University of Bern.

I am also very grateful to my Serbian friends in Lausanne who were always present at the important moments, especially to Jelena Mitić, Maria Blagojević, Alaksandra Stefanović, and Nadja Subotić. I will never forget my dear friend Saso Jezernik from the ETHZ. I am very thankful to my M.Sc.-thesis advisor, Prof. Dejan Popović, for his unconditional help and encouragement during my work towards the Ph.D. degree. Thanks to the other numerous friends who could not be listed here.

I would also like to thank William for his love and understanding. The greatest thanks go to my parents and to my brother. Without their support and love I would never become what I am today.

Slavica
Lausanne, December 2003

Contents

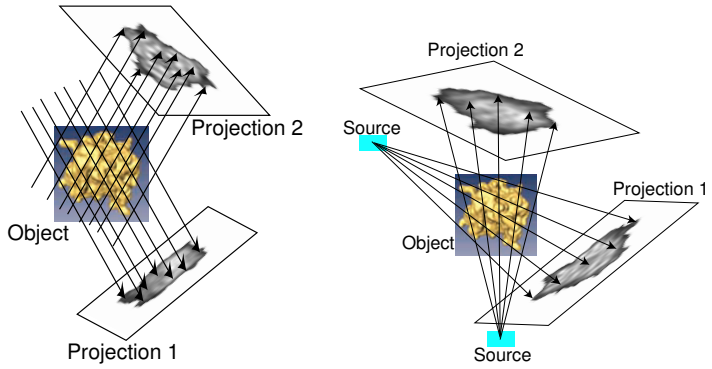
Chapter 1

INTRODUCTION

This thesis is about the alignment of a 3D model of an object to a set of 2D projections of the same object. The projections are acquired with a parallel- or a cone-beam projection geometry (Figure 1.1). The alignment problem can be formulated as follows: Compute the parameters of the volume pose at which the reference projection is as similar as possible to the projection that is simulated using the volume, for each of the given imaging views. In other words, the goal is to recover the transformation from the coordinate system attached to the 3D model of the object, that is, the volume coordinate system (V-COS), to the reference coordinate system (R-COS) (Figure 1.2) in which the reference projections have been acquired. This problem belongs to the group of 3D-to-2D registration problems. We will refer to it as volume-to-image registration.

1.1 What is registration?

The registration of two sets of points assumes two steps: 1) the determination of a geometric transformation between the sets, and 2) the application of the transformation to one of them, the test set, to align it with the other, the reference set. The most general transformation between the sets is elastic, which means that a straight line is mapped onto a curve. If the transformation preserves the distance between any two points, it is called a rigid-body transformation. The need for registration exists in many fields like medicine, biology, robotics, computer vision, pattern recognition, computer graphics, augmented-reality systems, manufacture [3, 9, 63, 64, 82, 90, 118].



(a) A parallel-beam projection is collected if the rays are parallel.

(b) A cone-beam projection is collected if all the rays meet in one location.

Figure 1.1: Projections are formed by measuring the total x-ray attenuations along straight lines through the object. (a) Parallel-beam projection geometry. (b) Cone-beam projection geometry.

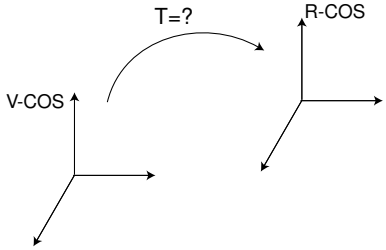


Figure 1.2: The goal of volume-to-image registration is to recover the transformation from the volume coordinate system (V-COS) to the reference coordinate system (R-COS) in which the reference projections have been collected such that the simulated projections are as similar as possible to the reference ones.

Figure 1.3: 3D-to-2D registration problem: Given the desk's shadow on the wall at your right, find the pose of the desk.

Image registration techniques align images of same or different dimensionality, the most often two-dimensional (2D) and three-dimensional (3D) data (e.g., 2D-to-2D, 3D-to-3D, 2D-to-3D, 3D-to-2D). These images may be acquired at different times, from different sensors, from different viewpoints, or from different objects [15, 39, 110]. This thesis is concerned with the techniques for the 3D-to-2D image registration.

1.2 Intuitive 3D-to-2D registration

It is a sunny day. We are in a room with a window and a desk is placed along a wall as in Figure 1.3. The desk makes a shadow on the wall. Let us now imagine that someone cleaning the room has moved the desk to some other place but its shadow on the wall has been captured by some means. Imagine that we have been assigned the following task: Return the desk to its original position by matching its actual shadow to the captured one. Obviously, to make it, we have to come the next day at the same time at which the shadow has been captured the day before, that is, when there is the same sunlight in the room. The positioning of the desk, by matching its reference and current shadows, is an illustration of what we are doing when performing the 3D-to-2D image registration. In this case, we would return the desk in a few simple movements, that is, we would only have to shift it and/or to turn it. In a more complicated case, for instance, of having to position a cat (instead of a desk) by shadow matching (Figure 1.4), we would have to account for more complex, free movements of the cat. In the case of the desk, we would therefore perform a rigid-body registration of the desk to its reference shadow while, in the case

(a) Cat at pose 1. (b) Cat at pose 2.

Figure 1.4: 3D-to-2D registration problem: Given a set of cat's shadows on the floor, find its pose for each shadow.

of the cat, we would have to perform an elastic registration of the cat to her reference shadow.

Less hypothetical is the following example of the 3D-to-2D registration problem: Given a set of pictures of a static toy, find the pose of the camera for each picture (Figure 1.5). To solve this problem, we can reformulate it as: Given a set of pictures of a static camera, find the pose of the toy for each picture (Figure 1.6). These two problems have the same solution. The principle of the solution is shown in Figure 1.7 and consists of matching the 2D reference and simulated silhouettes of the toy.

In this thesis, we will consider that we have access to more than just silhouettes. By replacing sunshine by x-rays, we will have access to the interior of the objects; this will make for richer, but more challenging data.

1.3 Applications of volume-to-image registration

The need for volume-to-image registration arises in many areas. Let us mention three of them: Computer-Assisted Orthopedic Surgery (CAOS), radiotherapy treatment, and 3D Electron Microscopy (3D EM).

CAOS [2, 24] assumes the integration of preoperative 3D Computed Tomography (CT) data [41, 67, 68] into the intraoperative procedure based on their registration [19, 20, 25, 34, 36, 51, 53, 74, 81, 95, 100, 117] to a set of intraoperatively acquired images of the patient (Figure 1.8).

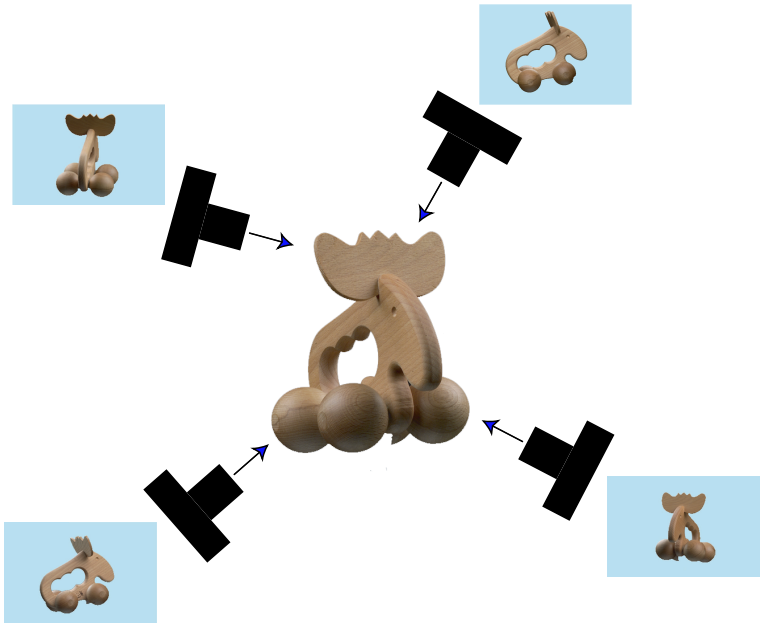


Figure 1.5: 3D-to-2D registration problem: Given a set of pictures of a static toy, find the pose of the camera for each picture.

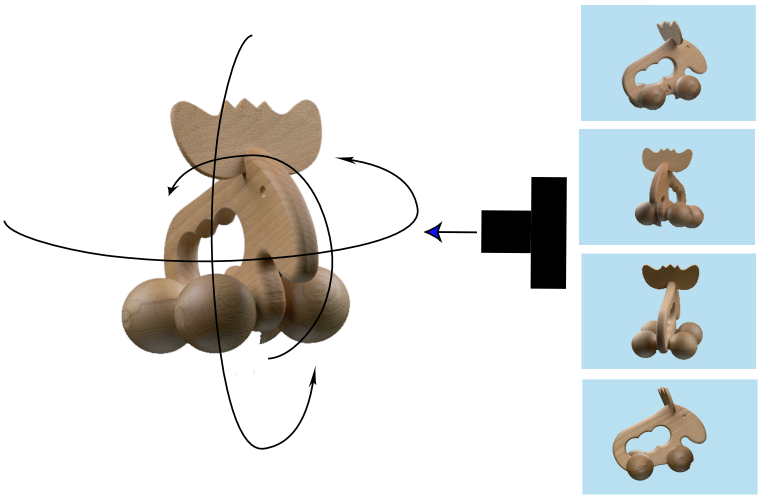
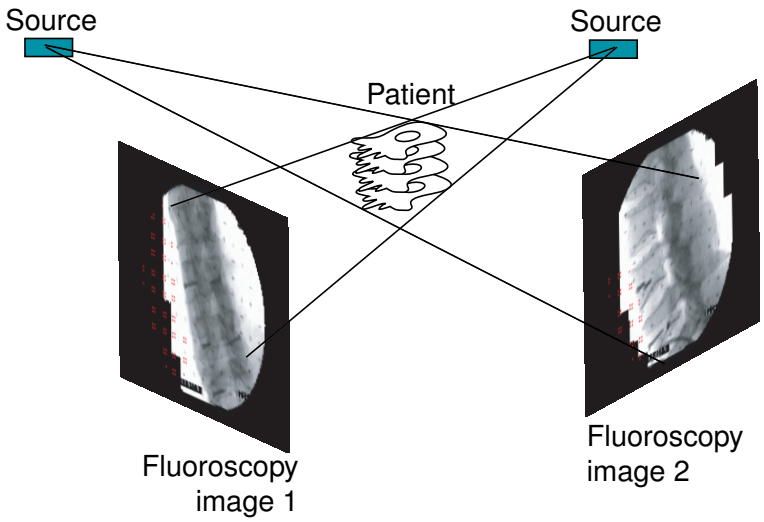
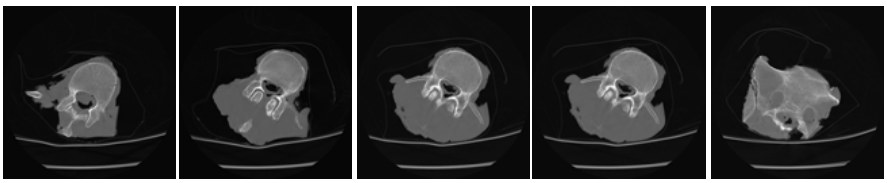


Figure 1.6: 3D-to-2D registration problem equivalent to the one shown in Figure 1.5: Given a set of pictures of a static camera, find the pose of the toy for each picture.

Figure 1.7: Solution of the registration problems from Figures 1.5 and 1.6 found by matching the toy's 2D reference and simulated silhouettes. Left: the reference silhouette. Middle: the simulated silhouette. Right: the difference between the reference and simulated silhouette.



(a) Physical space.



(b) 3D CT representation of the patient.

Figure 1.8: Application of the volume-to-image registration in image-guided orthopedic surgery. A 3D CT of the patient is aligned to its fluoroscopy (x-ray) images to display the surgical tools in the 3D CT during the intervention.

Radiotherapy treatment requires 3D pretreatment planning CT data to be registered to the prescribed position of the x-ray beam relative to the patient (Figure 1.9). This is usually done by registering 2D portal radiograph images acquired during the treatment to the 3D CT data [4, 16, 31, 54, 86, 91].

3D EM [26] yields a 3D reconstruction of a particle from the particle images (Figure 1.10) whose orientations are unknown and have to be estimated [21, 30, 32, 37, 52, 71, 73, 79, 93, 111] before applying an algorithm for 3D reconstruction. Given an earlier 3D reconstruction, the volume-to-image registration can be used to refine these estimates. A new 3D reconstruction based on a better estimate of the pose of the particle images should improve on the previous one.

The first two applications involve a cone-beam projection geometry while the third one is based on a parallel-beam projection geometry. In this thesis, we will show the potential of our registration algorithms to be used in two of the three areas: in orthopedic surgery and in 3D EM.

1.4 Algorithms for volume-to-image registration

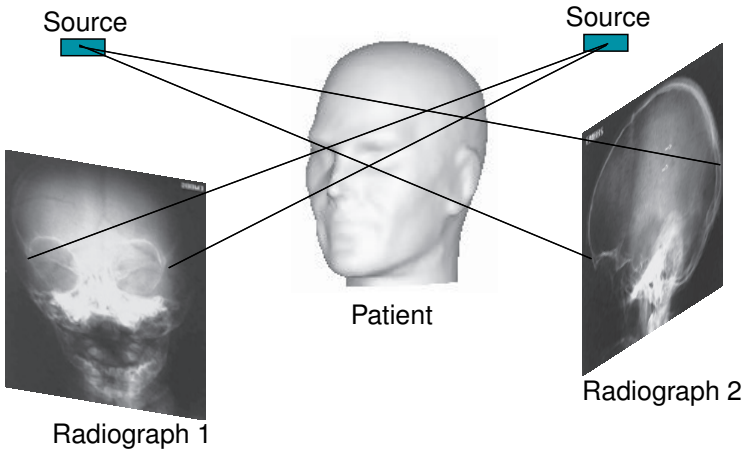
We can classify existing volume-to-image registration algorithms in two groups.

One group of algorithms is based on matching features that have first to be extracted from both the volume and the image. They match either markers (artificial features) implanted onto the object prior to data acquisition [2, 95], or anatomical features [19, 20, 25, 34, 35, 36, 53, 100]. The registration based on artificial markers is not enthusiastically accepted for medical applications because of its invasiveness.

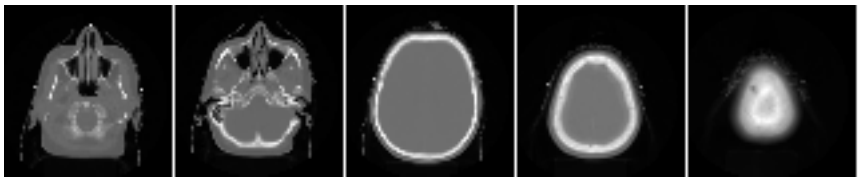
The registration based on anatomical features requires an accurate segmentation which is difficult to achieve fully automatically. It takes advantage of an extension of the standard fast Iterative Closest Point (ICP) algorithm for registration of curves and surfaces. However, the time required for accurate segmentation makes it unattractive for surgical interventions. An additional disadvantage is a low robustness against a partial data problem (presence of a feature in one imaging modality and absence in the other).

Neither of the two approaches can be applied for electron-microscopy data registration since the true 3D structure is not known in that application. Feature extraction makes therefore no sense there.

Projection-based registration methods form the other group [16, 21, 37, 51, 54, 74, 75, 81, 83, 86, 93, 116, 117]. In principle, all methods of this group achieve the registration by matching (Figure 1.11) a set of reference images to

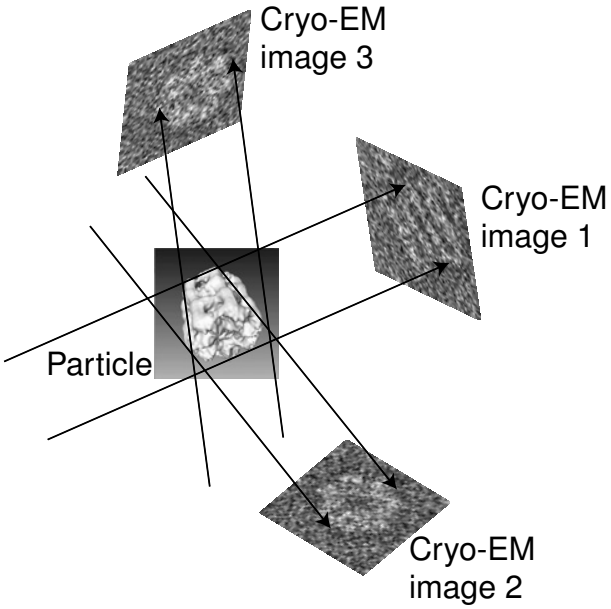


(a) Physical space.

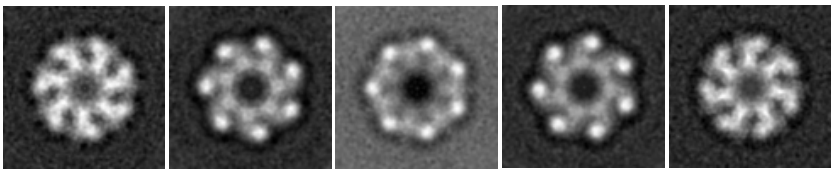


(b) 3D CT representation of the patient.

Figure 1.9: Application of the volume-to-image registration in radiotherapy. A 3D CT of the patient is registered to its radiograph (x-ray) images to compute the pose of the patient during the treatment.



(a) Physical space.



(b) 3D tomographic representation of the particle.

Figure 1.10: Application of the volume-to-image registration in 3D EM. A 3D model of the particle is registered to its Electron Cryo-Microscopy (Cryo-EM) images to refine their poses such that a more accurate particle model can be reconstructed.

their simulations. The simulations are traditionally computed by integrating the volume intensities along simulated rays through the volume (Figure 1.12). In general, these methods require little or no segmentation. The matching is done by optimizing a measure of similarity between the reference images and their simulations. If they are based on an iterative optimization procedure that uses an initial estimate for the unknown parameters, their robustness depends on the quality of the estimate. When significantly close to the solution initially, they promise a high accuracy because they are using a large amount of data. Their speed decreases with an increase in data size. Among all volume-to-image registration methods, the projection-based approach is most attractive for surgical interventions because of its non-invasiveness and accuracy. Since it does not require feature extraction, it can be used for electron-microscopy data registration as well.

This thesis proposes two novel projection-based methods for their applications in orthopedic surgery and electron microscopy.

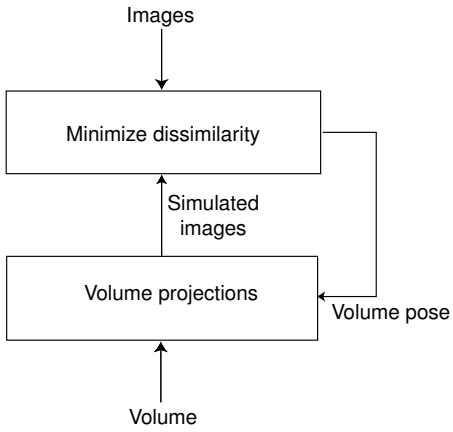


Figure 1.11: Diagram of the projection-based method for volume-to-image registration.

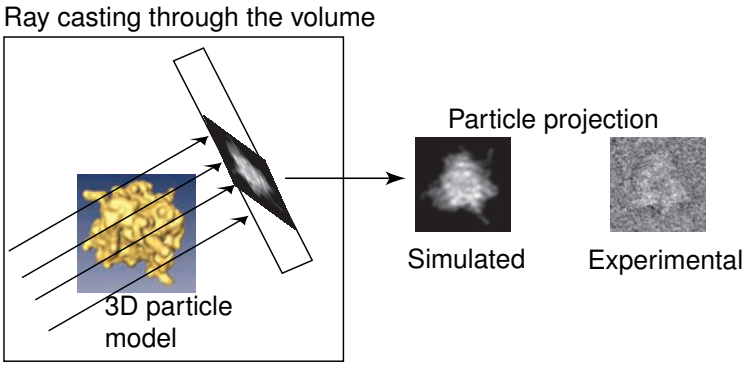


Figure 1.12: Ray casting. Left: Illustration of the method of simulating 2D projections of an object by casting the simulated (in this example, parallel-beam) rays through a 3D model of the object. Right: simulated and measured (experimental) projections of a protein.

1.5 Projection-based registration: State of the art

1.5.1 Geometric transformation

Projection-based registration in all three data cases, orthopedic, radiotherapy, and electron-microscopy, means the determination of the rigid-body transformation from the V-COS to the R-COS such that the simulated projections are as similar as possible to the reference ones [16, 21, 37, 51, 54, 74, 75, 81, 83, 86, 93, 116, 117].

1.5.2 Similarity measure

Penney et al. [75] have compared the accuracy of the registration of a 3D CT to a fluoroscopy image of a spine phantom when using one of six selected similarity measures. The ground-truth registration was determined by fiducial markers implanted on the specimen before the data acquisition. They reported pattern intensity [117] and gradient difference as being able to register accurately and robustly, even when soft tissues and interventional instruments were present in the fluoroscopy image. They found mutual information to be the least accurate of the six similarity measures.

Lemieux et al. [54] propose to register a 3D CT to two radiographic images

by optimizing a cross-correlation-based cost function at the beginning, and a gradient-based cost function at the final stage of the registration. With the same goal, Brown and Boult [16] maximize gradient correlation between the original and the simulated radiographs.

Similarly, the cross-correlation [73, 93] and the mean-squares dissimilarity measure [37] have been used to compute the similarity between the reference image and the volume projections for 3D EM.

The reference image and the projection of the volume may have different intensity ranges. Brown and Boult [16] explore the physical relationship between CT and radiograph measurements. They apply this relationship and correct the radiographs to improve the accuracy of their registration to the CT volume.

Penney et al. [75] propose to use a suitable intensity scaling factor as part of the similarity measure to adapt the intensity range of a CT-volume projection to that of a fluoroscopy image.

In this thesis, for simplicity purposes, we shall select the similarity measure to be the least-squares difference between the reference projections and their simulations after normalizing their respective intensity ranges.

1.5.3 Domain of the volume interpolation

Volume projections can be computed in two ways. One, which we mentioned already, is based on casting rays through the volume and requires interpolation of the volume in real space. The other is suited to the parallel-beam geometry and is based on the Central Slice Theorem (CST). It therefore requires interpolation in the transform domain. The CST results in the 2D transform of a projection of the volume being obtained by extracting the corresponding slice through the origin of the 3D transform of the volume. The orientation of the slice is such that it is perpendicular to the projection direction. The CST for the Fourier Transform (FT) is illustrated in Figure 1.13.

Real space

The projection of a volume by ray casting creates artificial edges and gradients due to the finite size of the volume, that is, due to the inequal lengths of the rays through such a volume. This problem does not occur in the real, collected images since the object that is imaged is treated as having an infinite extent. Lemieux et al. [54] and Penney et al. [75] propose to reduce the effect of the artificial edges by discarding some rays that produce them.

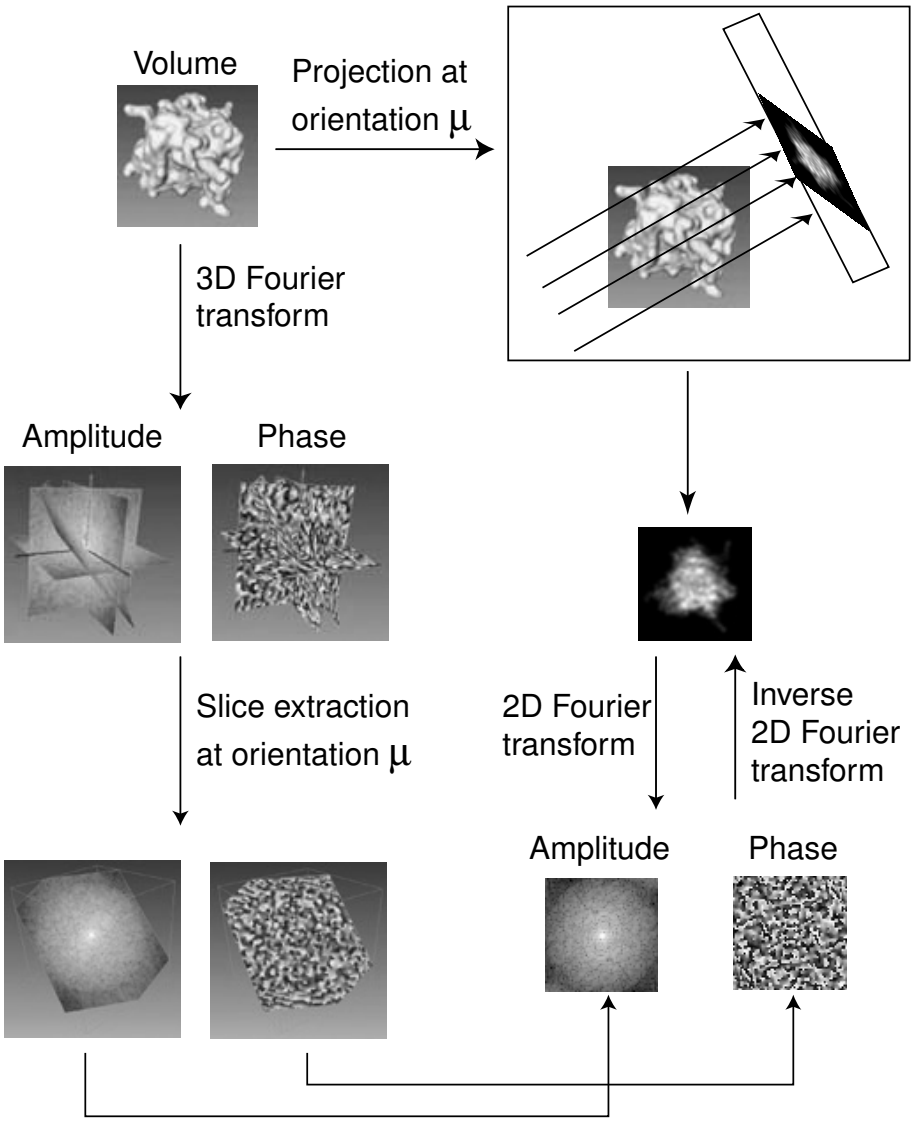


Figure 1.13: Central slice theorem for the Fourier transform.

Some methods have been proposed for the fast computation of the projections in real space [51, 83, 116]. They are based on ideas from computer graphics. The method called shear-warp factorization [116] consists of two steps. In a first step, an intermediate image is computed. The volume, resampled to a higher resolution, is projected using nearest-neighbor interpolation in-slice only. The second step relies on the linear interpolation of the intermediate image to obtain the final projection. Light fields, a technique to parameterize the set of all rays that emanate from a static scene, allow most of the computation to be performed in a preprocessing step [51, 83]. Using a light field generated for a set of volume projections, one can compute the other projections by interpolating the four-dimensional space of the prestored rays. These projections appear visually identical to those computed by the traditional ray casting [83].

In this thesis, we shall use ray casting to compute the volume projections for the space-based volume-to-image registration. We shall simulate ray-length equalization by projecting the volume from which the mean value has been subtracted. We shall use a novel one-step approach for fast ray casting called shearing that consists of replacing 3D interpolation of the volume by 2D interpolation.

Transform space

The CST has been used for high-speed volume rendering [18, 23, 96]. Dunne et al. [23] apply the CST for the FT (FT-CST) while Theussl et al. [96] implement the CST for the Hartley Transform (HT-CST). A family of algorithms for generating realistically shaded renderings of a volume using the FT-CST has been proposed in [55].

The FT-CST has been widely employed for reconstructing an object from its projection images [45, 48]. A reconstruction algorithm based on the HT-CST has been proposed in [66]. Imiya [44] uses the CST adapted to the exponential x-ray transform for the direct reconstruction of 3D images from line integrals of the emission type. Norton [70] benefits from the FT-CST for tomographic reconstruction of 2D vector fields with applications to flow imaging. Munson et al. [65] rely on the FT-CST to reconstruct images from the spotlight-mode synthetic-aperture radar that synthesizes high-resolution terrain maps using data gathered from multiple observation angles. The FT-CST is applied to reconstruct crosswell seismic direct-arrival traveltime tomograms in [80].

The CST has also been used for some other purposes. The FT-CST has been applied to synthesize a new x-ray image from a set of images exposed from a circular orbit above the patient that is identical in projection to a previous

x-ray image [17]. In 3D EM, some methods employ the FT-CST to determine the 3D pose of the particle images for a first 3D reconstruction of the particle. These methods [32, 52, 77, 111, 114] require no data interpolation.

In this thesis, we shall apply the FT-CST to compute the FTs of the volume projections for the frequency-based volume-to-image registration.

1.5.4 Interpolation kernel

Given some equidistant samples of a function, the ideal reconstruction kernel is the sinc function. However, this function is not acceptable in practice due to its slow decay. Thus, it is traditionally replaced by some approximation with faster decay. We show in Figure 1.14 some of these approximations. To compute the projections of a volume by ray casting, one usually interpolates the volume using linear interpolation [16, 54, 89]. To compute the projections by applying the FT-CST or the HT-CST, beside the linear interpolation [23], the Hamming windowed sinc function [96] and the technique of the Projection on Convex Sets (POCS) [33, 58] have been proposed.

POCS allows to create reconstruction functions that are truly limited in the frequency domain and that have a small amplitude outside some region in the space domain. It is based on an iterative procedure that chops off the tails of the transformation of the reconstruction function in the frequency domain and in the space domain, alternatively. This procedure starts with a Hamming-windowed sinc function in the frequency domain, iterates until the change from one iteration to the next is below some threshold, stopping with the truncation in the frequency domain.

Scaling functions of biorthogonal wavelets have been examined as reconstruction filters for frequency-domain volume rendering. It has been shown that their use reduces aliasing artifacts while maintaining efficiency [33].

In this thesis, we shall use splines for the volume interpolation in real space as well as in the domain of the FT, since these functions can lead to an analytically differentiable similarity measure. We shall select the spline degree such that we achieve a good tradeoff between the computational speed and interpolation quality.

1.5.5 Optimization algorithm

Different methods have been used to optimize the similarity measure such as the exhaustive-search technique [73, 93], Powell's multidimensional direction-set method [16, 54], and gradient-descent-type search techniques [74, 75].

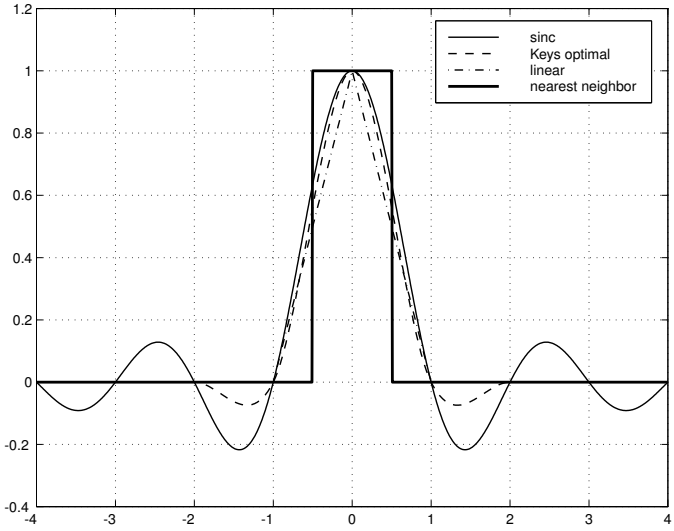


Figure 1.14: The sinc function and some of its finite-extent approximations (optimal Keys' cubic interpolant, the triangle function for the linear interpolation, and the rectangular pulse for the nearest-neighbor interpolation).

In this thesis, we shall employ a Levenberg-Marquardt (LM) optimizer since we have chosen to use an analytically differentiable least-squares similarity measure.

1.5.6 Size of the search space

A search within a continuous space of parameter values is commonly done in medical applications of volume-to-image registration [16, 54, 74, 75, 116, 117]. However, in 3D EM, a great number of images is usually processed. To reduce the time required for processing all particle images, one commonly does the search within a space containing discretized parameter values [73, 79, 93]. These methods are in general based on building a discrete set (library) of volume projections that are evenly distributed in space, and on matching the reference images to the library. An example of using a library of projections in medical applications is the method that has been proposed in [86] for registering a 3D CT to portal images. Another example is the iterative algorithm proposed in

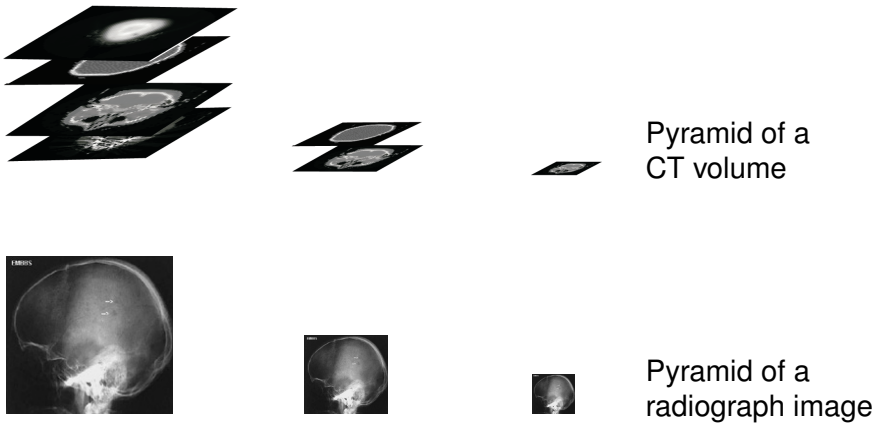


Figure 1.15: Multiresolution volume/image pyramids.

[37] which computes the projections of a volume at nine poses in the vicinity of the pose estimate from the previous iteration, and which determines the most similar projection to the reference image.

In this thesis, we shall deal with a continuous space of parameters in all cases since this yields a more accurate registration.

1.5.7 Multiresolution strategy

To improve the robustness against local optima of the similarity measure, the registration of a volume to a set of images is very often performed iteratively at multiple resolutions. Data of a lower resolution are usually obtained by blurring, using an averaging filter, and by their subsequent subsampling [54, 74, 75, 117].

In [54], a multiresolution pyramid of the volume was not computed explicitly but the projections of a lower resolution were computed by simply reducing the sampling rate along each ray through the volume, which introduces aliasing. A multiresolution pyramid of a region of interest extracted from the images and from the volume has been used in [75] and [117].

The method proposed in [93] computes the similarity between the reference images and the the library of volume projections based on their wavelet transform.

In this thesis, we shall use multiresolution volume/image pyramids (Figure 1.15) for the space-domain registration to make it more robust as well as to

speed it up. To improve the robustness of the frequency-domain algorithm, we shall benefit from a multiresolution, frequency-domain weighting of the similarity measure.

1.6 Contributions of the thesis: Summary

We develop here two novel algorithms for projection-based registration: space-based and frequency-based. Both algorithms are based on the LM iterative minimization of a least-squares dissimilarity between reference and simulated projections. The main novelty of our algorithms comes from the use of splines for volume interpolation in both the space domain and the frequency domain. We often use cubic splines. One advantage of using a cubic-spline data model is the possibility of having the gradient of the dissimilarity measure well-defined everywhere, which is a necessary condition for accurate registration.

We apply ray casting to compute the volume projections for the space-based registration. To speed up the computation, we introduce a novel one-step approach for fast ray casting that consists of replacing a 3D interpolation of the volume by a 2D interpolation. We reduce the effect of the unequal ray lengths by mean removal and variance normalization of the volume.

The frequency-based algorithm relies on the CST to compute the simulated projections in the FT domain for their comparison with the FTs of the reference projections.

To improve the robustness and the speed of the space-based algorithm, we take advantage of a coarse-to-fine processing of spline data pyramids. To improve the robustness of the frequency-based algorithm, we employ a frequency-domain weighting of the dissimilarity measure.

To achieve a more accurate registration, we deal with a continuous space of parameters in both the space-based approach and the frequency-based approach. This is a novelty in 3D electron-microscopy applications since the standard methods are based on a discretized space of parameters.

We have tested our algorithms for their applications in CAOS and in 3D EM. We have observed that the registration precision achieved using the orthopedic data was within the current standards. We have shown that our algorithm could be used in electron microscopy to refine the solution obtained by the standard algorithms.

1.7 Organization of the thesis

The data model is described in Chapter 2. The basic concepts of our spline projection-based registration in real space and in the frequency domain are presented in Chapter 3. The details of the space-based algorithm suited to the problem of registration of CT/C-arm orthopedic data are shown in Chapter 4, along with a validation study involving simulated data as well as real orthopedic data. The details of the frequency-based algorithm suited to the problem of registration of electron-microscopy data are given in Chapter 5, along with a validation study involving simulated data as well as real Electron Cryo-Microscopy (Cryo-EM) data. In Chapter 6, we compare the performance of the frequency-based algorithm to that of the space-based algorithm in the context of electron-microscopy registration problem. Chapter 7 contains conclusions and possible extensions.

Chapter 2

DATA MODELS

In this thesis, we have to interpolate volumetric data in the space domain as well as in the frequency domain. The theory of interpolation in real space and in the frequency domain is given in Section 2.1. First, we give a description of the traditional interpolation, although, we only use the generalized interpolation which is described in the second part of Section 2.1. More specifically, we use splines for interpolation; the corresponding data model is described in Section 2.2. For simplicity reasons, we consider only one-dimensional (1D) signals here. A generalization to more dimensions is straightforward and is presented in Section 2.2.6.

2.1 Interpolation and aliasing

2.1.1 Traditional interpolation

Real space

A continuously-defined function g can be reconstructed from the discrete samples $f_k = f(kT)$ of a function f (where T is the sampling step) as follows:

$$g(t) = \sum_{k \in \mathbb{Z}} f_k \, \varphi_{\text{int}}(\frac{t}{T} - k), \quad \forall t \in \mathbb{R}, \tag{2.1}$$

where φ_{int} is an interpolating reconstruction kernel which vanishes for all integer arguments except at the origin, where it has the value one [42, 97]. In this case, we have the desirable property that $g_k = f_k$.

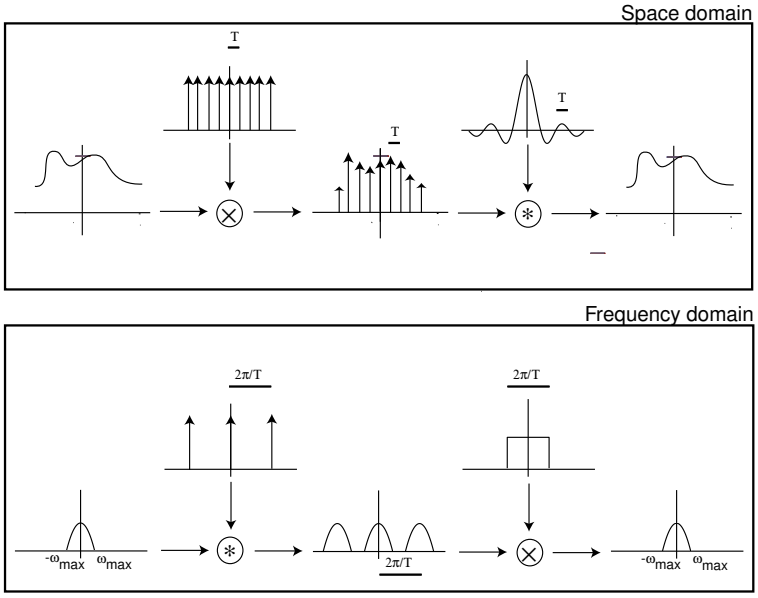


Figure 2.1: Sinc interpolation in the space domain that corresponds to the extraction of a single copy of the spectrum by looking through a rectangular window in the frequency domain.

Let us apply the FT to both sides of (2.1). We obtain

$$\hat{g}(\omega) = |T|\hat{\varphi}_{\text{int}}(\omega T) \sum_{k\in\mathbb{Z}} \hat{f}(\omega - \frac{2\pi k}{T}), \quad \forall \omega \in \mathbb{R}, \tag{2.2}$$

where $\hat{f}$, $\hat{g}$ are the FT of f, g, and where $\hat{\varphi}_{\text{int}}$ is the FT of φ_{int}. We accept without further discussion that these FTs exist and are well-defined.

On the right side of (2.2), the term multiplying $\hat{\varphi}_{\text{int}}(\omega T)$ is a $\frac{2\pi}{T}$-periodic function in ω. To avoid aliasing, we have to choose φ_{int} so that $\hat{\varphi}_{\text{int}}(\omega T)$ picks up only one copy of $\hat{f}$ from the replicas on the right. This is possible if two conditions are satisfied: 1) replicas do not overlap, and 2) $\hat{\varphi}_{\text{int}}(\omega T)$ has a support shorter than $\frac{2\pi}{T}$. To prevent the overlapping of the replicas, it is sufficient to limit the frequency content of the signal f to the range $(-\omega_{\max}, \omega_{\max})$, where

$\omega_{\max} \leq \frac{\pi}{T}$.

A classical example of the function φ_{int} is the sinc function, in which case we extract a single copy of the spectrum corresponding to $k = 0$ by looking through a rectangular window of length $\frac{2\pi}{T}$ (Figure 2.1). In this case, we achieve that $f = g$, a.e. (almost everywhere)

Frequency space

Let us now suppose that we have been given the samples $\hat{x}_n = \hat{x}(n\Omega)$ of the FT of a signal $x(t)$ where Ω is the sampling step in the frequency domain. We can reconstruct some continuously-defined FT of a signal $y(t)$ as follows:

$$\hat{y}(\omega) = \sum_{n\in\mathbb{Z}} \hat{x}_n \, \varphi_{\text{int}}(\frac{\omega}{\Omega} - n), \quad \forall\omega \in \mathbb{R}. \tag{2.3}$$

Let us apply the inverse FT to both sides of (2.3). We obtain

$$\mathcal{F}^{-1}\{\hat{y}\}(t) = |\Omega|\mathcal{F}^{-1}\{\varphi_{\text{int}}\}(t\Omega) \sum_{n\in\mathbb{Z}} \mathcal{F}^{-1}\{\hat{x}\}(t - \frac{2\pi n}{\Omega}), \quad \forall t \in \mathbb{R}, \tag{2.4}$$

where $\mathcal{F}^{-1}$ stands for the inverse Fourier transform.

We find, once more, that the reconstruction equation contains a term that is periodic with the period $\frac{2\pi}{\Omega}$. Therefore, the perfect reconstruction characterized by $y = x$, a.e., can be obtained when the two following conditions are satisfied: 1) $\mathcal{F}^{-1}\{\hat{x}\}$ is limited in space to the range $(-t_{\max}, t_{\max})$, where $t_{\max} \leq \frac{\pi}{\Omega}$, and 2) φ_{int} is selected such that $\mathcal{F}^{-1}\{\varphi_{\text{int}}\}(t\Omega)$ has a support shorter than $\frac{2\pi}{\Omega}$.

As in the space domain, a classical example of the function φ_{int} is the sinc function, in which case, we extract a single copy of $\mathcal{F}^{-1}\{\hat{x}\}$ corresponding to $n = 0$ by looking through a rectangular window of length $\frac{2\pi}{\Omega}$ (Figure 2.2). In this case, we achieve that $\hat{y} = \hat{x}$, a.e.

2.1.2 Generalized interpolation

Real space

The generalized interpolation of the signal samples f_k is performed as follows:

$$g(t) = \sum_{k\in\mathbb{Z}} c_k \, \varphi(\frac{t}{T} - k), \quad \forall t \in \mathbb{R}, \tag{2.5}$$

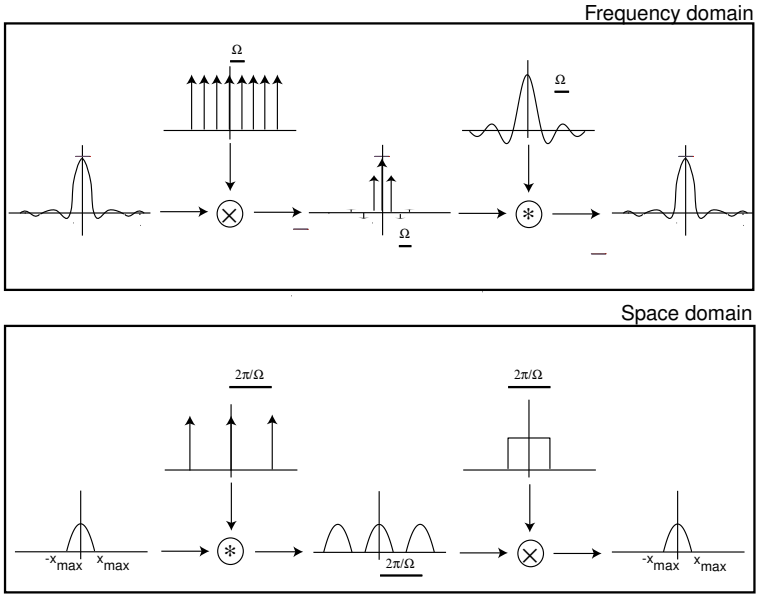


Figure 2.2: Sinc interpolation in the frequency domain that corresponds to the extraction of a single copy of the signal by looking through a rectangular window in the space domain.

where c_k are coefficients computed from the samples f_k [97] which we describe next. In general, $c_k \neq f_k$. The drawback of using the generalized formulation over the traditional formulation is adding one more step (the computation of c_k). This drawback is compensated by the gain in quality since this formulation offers an extended choice of basis functions φ that need not be interpolating anymore. Some of them have better properties than φ_{int} [97].

Coefficients via digital filtering. Given the signal samples f_k, we determine the coefficients c_k of the model (2.5) such that we have exact interpolation for $t = k_0 T$ [97], that is, we impose

$$g_{k_0} = f_{k_0} = \sum_{k \in \mathbb{Z}} c_k \, \varphi(k_0 - k), \forall k_0 \in \mathbb{Z}. \tag{2.6}$$

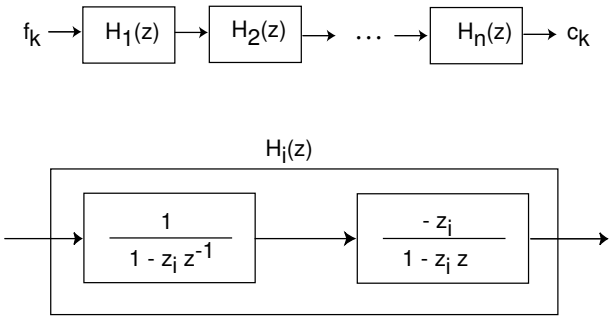


Figure 2.3: Cascade of first-order causal (with poles z_i) and anticausal (with poles z_i^{-1}) recursive IIR filters for the computation of the coefficients c_k used for the generalized interpolation of the signal samples f_k.

Let us introduce p as the discrete version of φ, that is, $p_k = \varphi(k), k \in \mathbb{Z}$. Then, we can rewrite (2.6) in the form of a discrete convolution

$$f_{k_0} = (c * p)_{k_0}. \tag{2.7}$$

The solution for c can be obtained as follows:

$$c_{k_0} = ((p)^{-1} * f)_{k_0}, \tag{2.8}$$

where $(p)^{-1}$ is the inverse convolution operator that generally does exist for the cases of interest. The convolution in (2.8) can be implemented by means of a digital filter. If we assume that φ is a symmetric function with finite support, which is generally the case in image processing, then we write the z-transform of $(p)^{-1}$ as

$$P^{-1}(z) = \frac{1}{\sum_{k\in\mathbb{Z}} p_k\, z^{-k}} \propto z^n \prod_{i=1}^{n} (z - z_i)^{-1} (z - z_i^{-1})^{-1}, \tag{2.9}$$

where $(2n + 1)$ is the number of integer indexes within the support of φ, and where $\{z_i\}$ are n out of $2n$ poles of $(p)^{-1}$ that are real and come in reciprocal pairs. From (2.9), it follows that the filter $(p)^{-1}$ can be implemented efficiently using a cascade of first-order causal (with the poles z_i) and anticausal (with the poles z_i^{-1}) recursive Infinite Impulse Response (IIR) filters [104, 105] (Figure 2.3).

Relationship between the traditional interpolation and the generalized interpolation. Let us combine (2.5) and (2.8) as

$$g(t) = \sum_{k_1 \in \mathbb{Z}} ((p)^{-1} * f)_{k_1} \, \varphi(\frac{t}{T} - k_1) \tag{2.10}$$

$$= \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} (p)^{-1}_{k_2} \, f_{k_1 - k_2} \, \varphi(\frac{t}{T} - k_1) \tag{2.11}$$

$$= \sum_{k_3 \in \mathbb{Z}} f_{k_3} \sum_{k_2 \in \mathbb{Z}} (p)^{-1}_{k_2} \, \varphi((\frac{t}{T} - k_2) - k_3). \tag{2.12}$$

Comparing (2.1) and (2.12), we identify the interpolating basis function φ_{int} hidden behind the noninterpolating φ as

$$\varphi_{\text{int}}(t) = \sum_{k \in \mathbb{Z}} (p)^{-1}_{k} \, \varphi(\frac{t}{T} - k). \tag{2.13}$$

We benefit from the generalized interpolation in the following sense. We use finite-support functions φ to efficiently perform the interpolation that relies on infinite-support functions φ_{int}.

Frequency space

The generalized interpolation of the FT samples $\hat{x}_n$ is performed as follows:

$$\hat{y}(\omega) = \sum_{n \in \mathbb{Z}} d_n \, \varphi(\frac{\omega}{\Omega} - n), \quad \forall \omega \in \mathbb{R}, \tag{2.14}$$

where d_n are coefficients obtained by digital filtering of the samples $\hat{x}_n$ using the same procedure as for the generalized interpolation in the space domain.

2.2 Spline data model

2.2.1 B-splines

The B-spline [88, 97, 102, 103, 104, 105] of degree n, β^n, is a center-symmetric, bell-shaped function (Figure 2.4) generated by the $(n+1)$-fold convolution of a rectangular pulse β^0

$$\beta^n(x) = \underbrace{\beta^0 * \beta^0 * \cdots * \beta^0}_{(n+1) \text{ times}}(x). \tag{2.15}$$

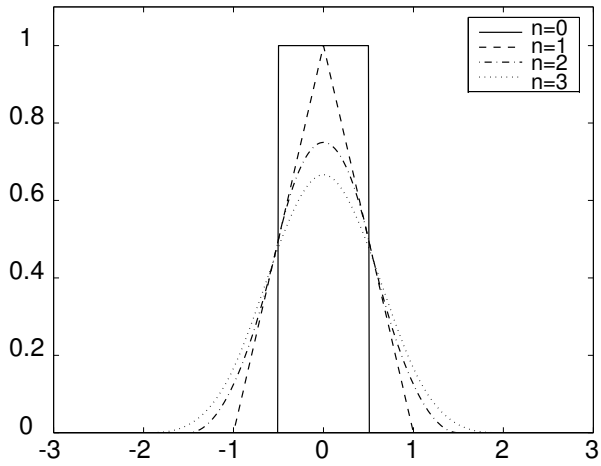


Figure 2.4: Centered B-splines of degree $n = 0, 1, 2,$ and 3.

An explicit formula for $\beta^n(x)$ is given by

$$\beta^n(x) = \sum_{k=0}^{n+1} \left(\begin{array}{c} n+1 \\ k \end{array} \right) (-1)^k \varsigma^n \left(x + \frac{n+1}{2} - k \right), \tag{2.16}$$

with $\varsigma^n(x) = \frac{1}{2n!} \operatorname{sign}(x)\, x^n$. Equation (2.16) shows that β^n is a piecewise polynomial of degree n that is continuously differentiable up to order $(n-1)$. Similarly, causal B-splines are given by

$$\beta_+^n(x) = \sum_{k=0}^{n+1} \left(\begin{array}{c} n+1 \\ k \end{array} \right) (-1)^k \varsigma^n \left(x - k \right). \tag{2.17}$$

For illustration purposes, we show in Figure 2.5 the spline interpolating kernels of degree $n = 1$ and 3 as well as the sinc function. In Figure 2.6, we show the frequency response of the spline interpolators of degree $n = 1$ and 3 compared to the frequency response of the ideal low-pass filter. It has been proved that the spline interpolators tend to the ideal low-pass filter as their degree increases [1].

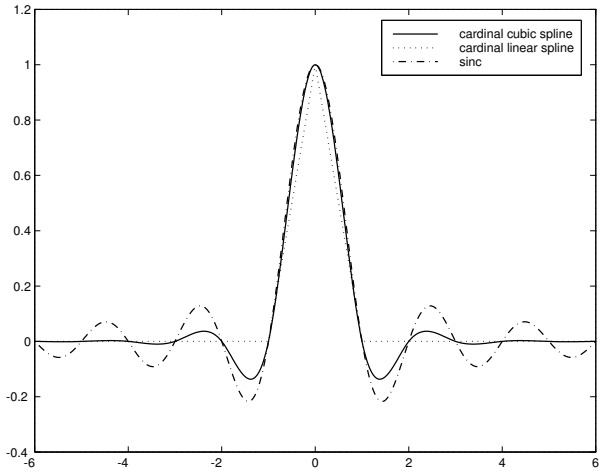


Figure 2.5: The spline interpolators of degree $n = 1$ (linear) and $n = 3$ (cubic) as well as the ideal sinc interpolator.

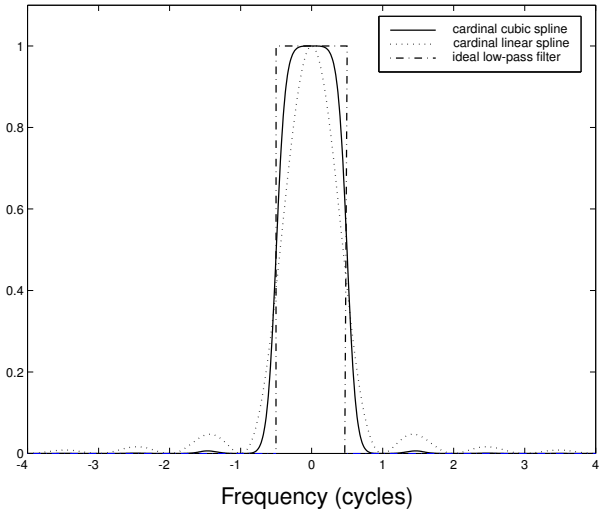


Figure 2.6: Frequency response of the spline interpolators of degree $n = 1$ (linear) and $n = 3$ (cubic) compared to that of the ideal low-pass filter.

2.2.2 Order of approximation

Let f_T be the function that has been reconstructed from the samples of a function f, that is, $f_T(x) = \sum_{k\in\mathbb{Z}} c_k\, \varphi(\frac{x}{T} - k)$. We can define the error of approximation of f by f_T using the mean-square norm: $\varepsilon^2(T) = \|f - f_T\|^2_{L_2} = \int_{-\infty}^{\infty} (f(x) - f_T(x))^2\, \mathrm{d}x$. A prediction of this approximation error can be obtained using the formula: $\eta^2(T) = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 E(\omega T)\, \mathrm{d}\omega$ [11, 12, 13]. The interpolation error kernel E depends on the basis function only. For bandlimited functions, we have that ε is equal to η, and the prediction is exact in this case.

The kernel E must vanish at the origin to ensure that the approximation error vanishes for $T = 0$. The vanishing rate is controlled by two properties of the basis function φ: the approximation order L and the approximation constant $C > 0$. They determine η as follows:

$$\eta(T) = C\, T^L\, \|f^{(L)}\|_{L_2} \qquad \text{as} \quad T \to 0, \tag{2.18}$$

where $f^{(L)}$ is the L-th derivative of f. This means that the error predicted by η decreases like T^L as $T \to 0$. The higher is L, the faster is the decay of the approximation error for the same, sufficiently small T. Also, the smaller is C, the smaller is the approximation error. However, the decay of the approximation error is dominated by T^L for $T \leq C$ which is very often the case. Therefore, C is used for ranking the basis functions of identical approximation order L [97]. See [97] for a systematic overview of standard basis functions.

It turns out that the approximation order (and the support) of a B-spline of degree n is equal to $L = n + 1$ [97]. The B-splines are shown to be maximally continuous basis functions, with the minimal support for a given order of approximation, and with the maximal order of approximation for a given support [10]. Cubic B-splines ($n = 3$) offer a good tradeoff between the computational cost and the interpolation quality. They are the functions that we are most often using.

2.2.3 m-scale relation

An interesting property of the causal B-splines of degree n is the m-scale relation [102]

$$\beta_+^n(x/m) = \sum_{k\in\mathbb{Z}} h_m(k)\, \beta_+^n(x - k), \tag{2.19}$$

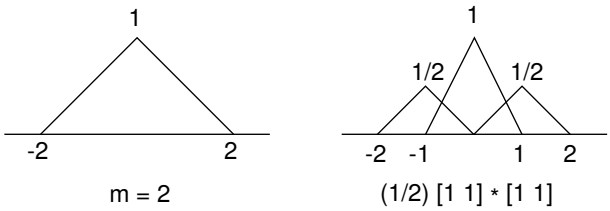


Figure 2.7: Two-scale relation for the centered B-spline of degree 1. From [102].

which says that a causal B-spline of degree n dilated by m is a linear combination of causal B-splines of same degree, with discrete weights $h_m(k)$ whose z-transform is

$$H_m(z) = \frac{1}{m^n} \left(\sum_{k=0}^{m-1} z^{-k} \right)^{n+1} . \tag{2.20}$$

This relationship holds for any positive integer m. A similar m-scale relation exists for centered B-splines if n and m are not both even [107]. If they are both even, additional shifts of the B-spline must be considered [99]. The m-scale relation of B-splines makes a basis for building spline multiresolution data pyramids [14, 106] that are in a close connection with spline wavelets [107]. An illustration of the two-scale relation for centered linear B-splines is given in Figure 2.7. We take advantage of the m-scale relation for multiresolution in the space domain.

2.2.4 First derivative of the B-splines

The B-splines are easy to manipulate and one of their interesting properties is that their derivatives can be computed recursively [102], starting with

$$\frac{\mathrm{d}}{\mathrm{d}x}\beta^n(x) = \beta^{n-1}\left(x + \frac{1}{2}\right) - \beta^{n-1}\left(x - \frac{1}{2}\right). \tag{2.21}$$

For $n \geq 2$, we have the guarantee that $\frac{\mathrm{d}}{\mathrm{d}x}\beta^n(x)$ is continuous, which is not the case if $n = 0$ or $n = 1$. The first derivatives of the centered B-splines of degree $n = 0, 1, 2,$ and 3 are shown in Figure 2.8. We make use of the first derivative of the B-splines while optimizing the similarity measure. We use cubic B-splines since they offer a good tradeoff between the continuity of their first derivatives and the computation time.

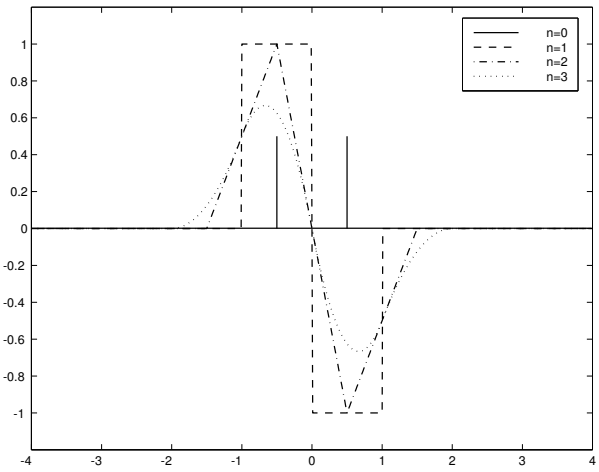


Figure 2.8: The first derivatives of the centered B-splines of degree $n = 0, 1, 2, \text{and } 3$.

2.2.5 Separability

A D-dimensional B-spline of degree n is defined as the separable function

$$\beta^n(\mathbf{x}) = \prod_{j=1}^{D} \beta^n(x_j), \quad \forall \mathbf{x} = (x_1, x_2, \ldots, x_D) \in \mathbb{R}^D. \tag{2.22}$$

This separability property makes possible some operations on D-dimensional data to be performed by a successive processing of 1D data along each of the D dimensions. In return, the data processing is simple and fast. We take advantage of the B-spline separability for a faster interpolation of our 2D and 3D data.

2.2.6 Model

Assuming that $T = 1$, the space-domain model (2.5), extended to D-dimensions and based on B-splines, is

$$g(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}^D} c_{\mathbf{k}} \, \beta^n(\mathbf{x} - \mathbf{k}), \quad \forall \mathbf{x} \in \mathbb{R}^D. \tag{2.23}$$

Similarly, assuming that $\Omega = 1$, the frequency-domain model (2.14), extended to D-dimensions and based on B-splines, is

$$\hat{y}(\boldsymbol{\omega}) = \sum_{\mathbf{n}\in\mathbb{Z}^D} d_{\mathbf{n}}\, \beta^n(\boldsymbol{\omega} - \mathbf{n}), \quad \forall \boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_D) \in \mathbb{R}^D. \tag{2.24}$$

We use the model (2.23) to interpolate data samples in the space domain, while the model (2.24) is applied to interpolate data samples in the frequency domain.

2.2.7 First derivative of the model

We write the gradient of the model in the space domain as

$$\boldsymbol{\nabla} g(\mathbf{x}) = \sum_{\mathbf{k}\in\mathbb{Z}^D} c_{\mathbf{k}}\, \boldsymbol{\nabla}\beta^n(\mathbf{x} - \mathbf{k}), \quad \forall \mathbf{x} \in \mathbb{R}^D. \tag{2.25}$$

The gradient of the model in the frequency domain is

$$\boldsymbol{\nabla} \hat{y}(\boldsymbol{\omega}) = \sum_{\mathbf{n}\in\mathbb{Z}^D} d_{\mathbf{n}}\, \boldsymbol{\nabla}\beta^n(\boldsymbol{\omega} - \mathbf{n}), \quad \forall \boldsymbol{\omega} \in \mathbb{R}^D. \tag{2.26}$$

Our space-domain optimization of the similarity measure relies on the gradient (2.25), while our frequency-domain optimization makes use of the gradient (2.26).

2.3 Spline data pyramids

In the multiresolution theory, one usually considers scaling factors that are powers of two, that is, $m = 2^i$.

Let us suppose that $\varphi(x/2^i - k)$ are the basis functions at the scale 2^i (φ enlarged by 2^i and spaced accordingly), and that $V_{2^i} = \text{span}_k\{\varphi(x/2^i - k)\}$ is the corresponding subspace of the space L_2. Let $f_i \in V_{2^i}$ be the minimum least-squares error approximation of a signal $f \in L_2$ at the scale 2^i

$$f_i(x) = \sum_{k\in\mathbb{Z}} c_{2^i}(k)\, \varphi(x/2^i - k). \tag{2.27}$$

The m-scale relation for $m = 2$ implies the nestedness of the subspaces: $V_1 \supset V_2 \supset \cdots \supset V_{2^i} \ldots$. Thanks to this property, the coefficients $c_{2^i}(k)$ can be

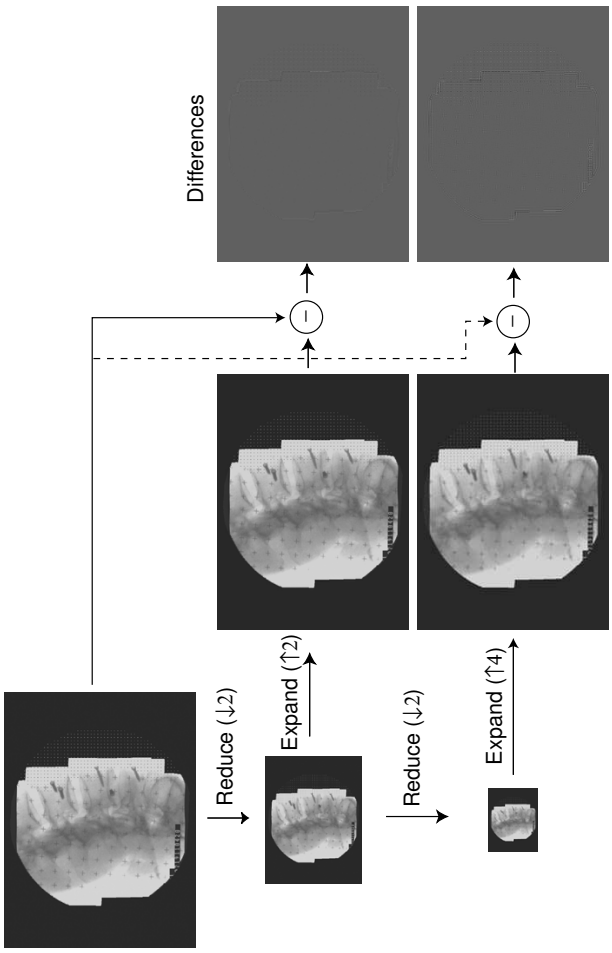


Figure 2.9: A two-level cubic-spline L_2 C-arm image pyramid. See the text for details.

computed iteratively using a prefilter $\mathring{h}(k)$ and a down-sampling by 2 [102]

$$c_{2^i}(k) = (\mathring{h} * c_{2^{i-1}})(2k). \tag{2.28}$$

Using the two-scale relation, the filter $\mathring{h}(k)$ can be expressed in the inner-product

form

$$\mathring{h}(k) = \langle \varphi(\cdot + k), \mathring{\varphi}(\cdot/2) \rangle, \tag{2.29}$$

where $\mathring{\varphi}$ is the dual of φ, which means that $\mathring{\varphi}$ satisfies the bi-orthogonality condition $\langle \mathring{\varphi}(\cdot - k), \varphi(\cdot - l) \rangle = \delta_{k-l}$, where δ is the Kronecker delta. In case of multidimensional data, the coefficients $c_{2^i}(k)$ are obtained by successive 1D filtering and decimation of the data along each of the dimensions.

A two-level L_2 cubic-spline image pyramid is shown in Figure 2.9. For illustration, the C-arm image was reduced twice and both coarse versions expanded back to the original image resolution. At the end, these expanded images were subtracted from the original image. The difference images are shown as well; the fact that they are essentially featureless is a testimony to the quality of the pyramid.

We use L_2 spline volume and image pyramids for space-domain multiresolution. Again, we create the pyramids based on cubic splines since they offer a good tradeoff between the computation time and approximation quality.

Chapter 3

SPLINE PROJECTION-BASED REGISTRATION: BASIC CONCEPTS

In this chapter, we describe at the same time our space-based and frequency-based algorithms for volume-to-image registration.

We estimate the pose $\boldsymbol{\mu}$ of the volume in the reference coordinate system (R-COS) with respect to the reference projections by refining an initial volume pose. Figure 3.1 shows a simplified block-diagram of both the space-based approach and the frequency-based approach. The input data for the algorithms are: 1) the volume, 2) the reference projections, 3) the projection-model geometrical data, and 4) the initial volume pose. The algorithms achieve the registration by iterative minimization of the dissimilarity between the reference projections and the simulated projections or between their respective FTs.

Error-free (sinc-based) interpolation in the FT domain can be achieved if the data are space-limited (Section 2.1.1), which is not always the case. To interpolate the data in the FT domain accurately enough, we assume that they are nearly space-limited. To enforce this assumption, we apodize the input volume and images by Gaussian windows with the appropriate standard deviations which should be selected by the user. In principle, they should not be smaller

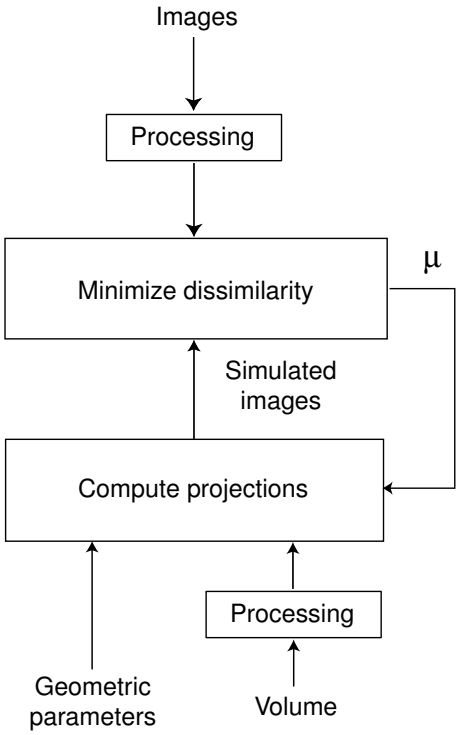


Figure 3.1: Our algorithm for volume-to-image registration.

than half of the object size. Note that the "Processing" blocks in Figure 3.1 stand for the Gaussian-window data apodization followed by the computation of the FTs in the case of the frequency-based registration. In the case of the space-based registration, these blocks perform the normalization of data intensities.

For simplicity, we describe the two registration algorithms using a parallel-beam projection geometry. We extend these ideas to a cone-beam geometry in Chapter 4.

We start this chapter with the definition of the volume pose, which can be defined in many ways. Here, we use a parameterization that is standard in 3D EM (Section 3.1). The optimization problem is described in Section 3.2. The least-squares criterion to optimize is given in Section 3.3. The ray casting used to compute the volume projections, as well as the CST-based method used to

compute the FTs of these projections, are explained in Sections 3.4 and 3.5, respectively. The multiresolution strategies for space-domain and frequency-domain registrations are investigated in Section 3.6. The LM algorithm, a gradient-based optimizer, is presented in Section 3.7.

3.1 Volume pose

We define the volume pose $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_6)$ in the R-COS by three rotations and three translations of the volume. Let us denote the system of 3D voxel indexes $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{R}^3$ by the voxel coordinate system (V-COS) and the system of 2D pixel indexes $\mathbf{k} = (k_1, k_2, 1) \in \mathbb{R}^2 \times \{1\}$ by the pixel coordinate system (P-COS). We rotate the volume by applying a 3×3 rotation matrix R to the V-COS coordinate of each voxel of the volume. This matrix is determined by three Euler angles: φ, θ, and ψ, that is, $R = R_z(\psi)R_y(\theta)R_z(\varphi)$, where $R_x(\alpha)$ indicates the matrix of rotation around the x-axis by the angle α. We translate the volume by shifting the V-COS coordinate of each voxel of the volume by a vector $\mathbf{t} = (\Delta x, \Delta y, \Delta z)$, where Δx, Δy, and Δz are translations along the x-, y-, and z-axes. We transform the V-COS coordinate of a point of the volume, $\mathbf{n}$, into its corresponding P-COS coordinate, $\mathbf{k}$, as $\mathbf{k} = H\Lambda(R\mathbf{n} + \mathbf{t}) + (0, 0, 1)$, where Λ is a scaling 3×3 diagonal matrix that accounts for a possible difference between the physical sizes of the pixels and the voxels, and H projects a 3D vector onto the xy-plane, that is, $H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Note that the transposed matrix is $H^\top = H$. Similarly, we transform the P-COS coordinate of a point inside the image, $\mathbf{k}$, into a line of corresponding V-COS coordinates, $\mathbf{n}_0 + \xi\mathbf{n}$, where $\xi \in \mathbb{R}$ is some free scalar parameter, with $\mathbf{n}_0 = R^{-1}(\Lambda^{-1}\mathbf{k} - \mathbf{t})$ and $\mathbf{n} = R^{-1}\Lambda^{-1}(0, 0, 1)$.

3.2 Optimization problem

We compute the pose of a volume $\boldsymbol{\mu}$ with respect to a set of its reference projections by minimizing a real-valued function S of N variables $\mu_i, i = 1, 2, ..., N$ that we refer to as the cost function. Let us write the solution of this problem as

$$\boldsymbol{\mu}^* = \arg\min_{\boldsymbol{\mu}} S(\boldsymbol{\mu}). \tag{3.1}$$

We say that this problem is unconstrained since we impose no conditions on $\boldsymbol{\mu}$ and assume that S is defined for all $\boldsymbol{\mu}$. $\boldsymbol{\mu}^*$ can be a local minimizer ($S(\boldsymbol{\mu}^*) \leq S(\boldsymbol{\mu})$ for all $\boldsymbol{\mu}$ near $\boldsymbol{\mu}^*$), or a global minimizer ($S(\boldsymbol{\mu}^*) \leq S(\boldsymbol{\mu})$ for all $\boldsymbol{\mu}$) [49].

Let us consider the case of a twice-continuously differentiable cost function. The necessary conditions for optimality imply that the gradient $\boldsymbol{\nabla} S(\boldsymbol{\mu}) = \left(\frac{\partial S}{\partial \mu_1}, \frac{\partial S}{\partial \mu_2}, \ldots, \frac{\partial S}{\partial \mu_N}\right)$ vanishes and that the Hessian

$$\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}) = \begin{pmatrix} \frac{\partial^2 S}{\partial \mu_1 \, \partial \mu_1} & \frac{\partial^2 S}{\partial \mu_1 \, \partial \mu_2} & \cdots & \frac{\partial^2 S}{\partial \mu_1 \, \partial \mu_N} \\ \frac{\partial^2 S}{\partial \mu_2 \, \partial \mu_1} & \frac{\partial^2 S}{\partial \mu_2 \, \partial \mu_2} & \cdots & \frac{\partial^2 S}{\partial \mu_2 \, \partial \mu_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 S}{\partial \mu_N \, \partial \mu_1} & \frac{\partial^2 S}{\partial \mu_N \, \partial \mu_2} & \cdots & \frac{\partial^2 S}{\partial \mu_N \, \partial \mu_N} \end{pmatrix} \tag{3.2}$$

is positive semidefinite at a local minimizer.

3.3 Cost function

The registration process will be driven by the gradient of the cost function. Fortunately, the spline data model makes the exact computation of the gradient possible. We select a least-squares cost function since it results in a simple expression for the gradient. Besides, least-squares criterions are known to be good for data fitting (parameter identification), which is our case since we choose to match the reference projections (or their FTs) to their respective models. In particular, least-squares are optimal if the data are corrupted by white Gaussian noise.

Let the reference projection in either of the two domains, the space domain or the FT domain, be given in a vector form. Let us denote this M-dimensional vector by $\mathbf{p}$, where M is the number of pixels in the reference projection. Let us suppose that we have simulated the projection of the volume at pose $\boldsymbol{\mu}$, which we denote by $\mathbf{p}_{\boldsymbol{\mu}} \in \mathbb{R}^M$. Then, we can compute their dissimilarity as $\mathbf{r}(\boldsymbol{\mu}) = \mathbf{p}_{\boldsymbol{\mu}} - \mathbf{p} = (r_1, r_2, \ldots, r_M)$. We propose to solve the registration problem as a nonlinear least-squares problem, that is, to minimize the cost function

$$S(\boldsymbol{\mu}) = \frac{1}{2} \sum_{q=1}^{Q} (\mathbf{r}_q(\boldsymbol{\mu}))^H \mathbf{r}_q(\boldsymbol{\mu}), \tag{3.3}$$

where $\mathbf{r}_q$ is the dissimilarity for the q-th imaging view, Q is the number of given imaging views, and where $\mathbf{r}^H$ denotes the Hermitian transpose of $\mathbf{r}$.

In our case, we have that $Q\,M \gg N$ (the problem is overdetermined), which hints at a robust solution to our registration problem.

3.4 Ray casting

In this section, we describe the method that computes a projection of the volume by simulating x-rays through the volume.

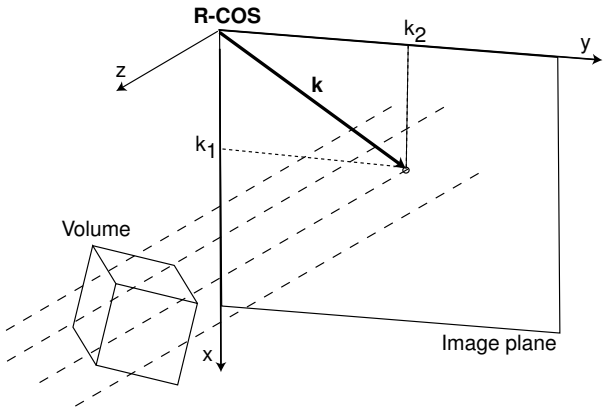


Figure 3.2: Parallel x-ray beam geometry.

Given a volume f at pose $\boldsymbol{\mu}$ in the R-COS, we cast the simulated rays parallel to the z-axis through each pixel of the image plane (Figure 3.2). We project the volume along the ray determined by the unit vector $\mathbf{e}_3 = (0, 0, 1)$ and the P-COS coordinate of a pixel $\mathbf{k} = (k_1, k_2, 1)$ as follows:

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \int_{\mathbb{R}} f\left(R^{-1}(\Lambda^{-1}\mathbf{k} - \mathbf{t}) + \xi\, R^{-1}\Lambda^{-1}\mathbf{e}_3\right)\, \mathrm{d}\xi. \tag{3.4}$$

We rewrite (3.4) as

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \int_{\mathbb{R}} f\left(\mathbf{n}_0 + \xi\, \mathbf{n}\right)\, \mathrm{d}\xi, \tag{3.5}$$

where $\mathbf{n}_0 = R^{-1}(\Lambda^{-1}\mathbf{k} - \mathbf{t})$ and $\mathbf{n} = R^{-1}\Lambda^{-1}\mathbf{e}_3$.

3.4.1 Shears

A direct implementation of (3.5) would require a 3D interpolation of the volume f at the V-COS coordinate $\mathbf{n}_0 + \xi\,\mathbf{n}$. We perform a fast computation of the projection by replacing the 3D interpolation by a 2D interpolation. We refer to this method as shearing [46]. The shearing is based on one out of three well-chosen changes of the integration variable ξ in (3.5). The new integration variable is selected from one of the following three possibilities $\xi_j, j \in \{1,2,3\}$

$$\xi_j = [\mathbf{n}_0]_j + \xi\,[\mathbf{n}]_j \longrightarrow \mathrm{d}\xi_j = [\mathbf{n}]_j\,\mathrm{d}\xi, \tag{3.6}$$

where $[\mathbf{x}]_j$ is the j-th component of the vector $\mathbf{x}$.

After expressing ξ and $\mathrm{d}\xi$ from (3.6) and their replacement into (3.5), this equation becomes

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \frac{1}{|[\mathbf{n}]_j|}\int_{\mathbb{R}} f\left(\mathbf{n}_0 - \frac{[\mathbf{n}_0]_j}{[\mathbf{n}]_j}\mathbf{n} + \xi_j\,\frac{1}{[\mathbf{n}]_j}\mathbf{n}\right)\mathrm{d}\xi_j,\ j \in \{1,2,3\}, \tag{3.7}$$

which we can condense into

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \lambda_j \int_{\mathbb{R}} f\left(\mathbf{n}_0^j + \xi_j\,\mathbf{n}^j\right)\mathrm{d}\xi_j,\ j \in \{1,2,3\}, \tag{3.8}$$

where $\lambda_j = \frac{1}{|[\mathbf{n}]_j|}$, $\mathbf{n}_0^j = \mathbf{n}_0 - \frac{[\mathbf{n}_0]_j}{[\mathbf{n}]_j}\mathbf{n}$, and $\mathbf{n}^j = \frac{1}{[\mathbf{n}]_j}\mathbf{n}$.

The scaling factor λ_j determines which one of the three expressions for $p_{\boldsymbol{\mu}}$ will be used to compute the projection. We chose that j for which λ_j is minimum. Since λ_j does not depend on $\mathbf{k}$, the decision about j is taken once for all pixels. An example of applying the shearing method for computing a 1D parallel-beam projection of an image is shown in Figure 3.3. The shearing method adapted to the cone-beam projection geometry is described in Chapter 4.

The phase adaptation proceeds by remarking that we can choose any finite $\Delta\xi_j$ and still write (3.8) as

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \lambda_j \int_{\mathbb{R}} f\left(\mathbf{n}_0^j + (\xi_j + \Delta\xi_j)\,\mathbf{n}^j\right)\mathrm{d}\xi_j,\ j \in \{1,2,3\}. \tag{3.9}$$

We choose $\Delta\xi_j$ such that $\left[\mathbf{n}_0^j + \Delta\xi_j\,\mathbf{n}^j\right]_j = 0$, which leads to a discrete version of (3.8) in which we approximate the integral by a discrete sum and adapt the

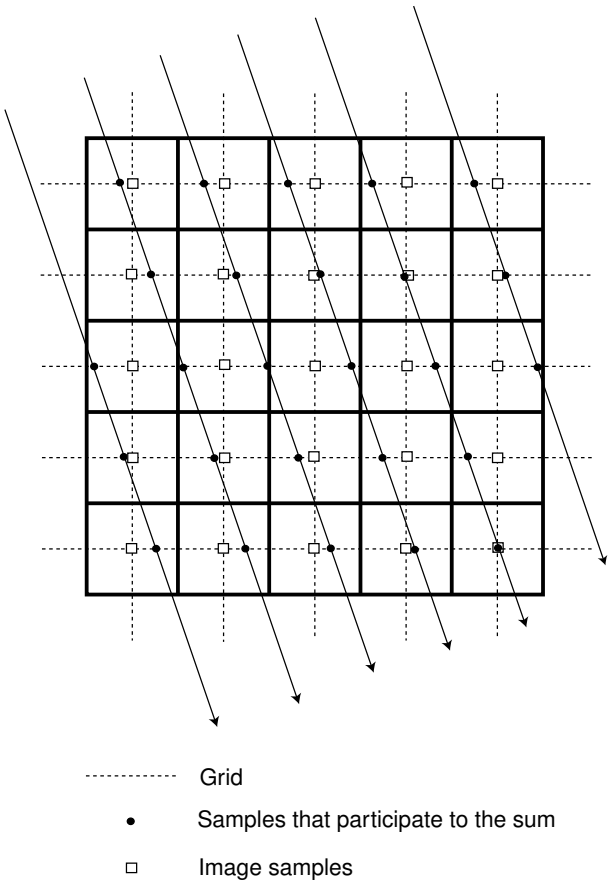


Figure 3.3: Shearing approach to the 1D parallel-beam projection of an image.

size and the phase of the sampling step along ξ_j so that only samples of the volume with one integer and two real coordinates take part in the sum, that is,

$$p_{\boldsymbol{\mu}}(\mathbf{k}) \approx \lambda_j \sum_{k\in\mathbb{Z}} f\left(\mathbf{n}_0^j + (k+\Delta\xi_j)\ \mathbf{n}^j\right),\ j \in \{1,2,3\}. \tag{3.10}$$

Note that, by construction, the evaluation of f requires the interpolation of the volume at a point with one integer and only two real coordinates. To interpolate the data, we use the B-spline data model (2.23).

3.4.2 Approximation of the integral by a discrete sum

Let us write the Poisson's summation formula for a function f

$$T\sum_{k\in\mathbb{Z}} f(kT) = \sum_{m\in\mathbb{Z}} \hat{f}\left(\frac{2\pi}{T}m\right),\ T>0. \tag{3.11}$$

Recalling that $\hat{f}(0) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$, we can rewrite (3.11) as

$$T\sum_{k\in\mathbb{Z}} f(kT) = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x + \sum_{m\in\mathbb{Z}\setminus\{0\}} \hat{f}\left(\frac{2\pi}{T}m\right). \tag{3.12}$$

Let ϵ denote the error of the approximation of the integral by a discrete sum

$$\epsilon = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x - T\sum_{k\in\mathbb{Z}} f(kT) = -\sum_{m\in\mathbb{Z}\setminus\{0\}} \hat{f}\left(\frac{2\pi}{T}m\right). \tag{3.13}$$

If the function f is bandlimited, we have that $\hat{f}(\omega) = 0$ for $|\omega| > \frac{\pi}{T}$, which implies that $\epsilon = 0$. If f is not bandlimited, but can be approximated by a bandlimited function, for instance one that is based on splines of degree n and that satisfies $|\hat{f}(\omega)| \leq \frac{C}{|\omega|^{1+n}}$, we have that

$$|\epsilon| \leq \sum_{m\in\mathbb{Z}\setminus\{0\}} \frac{C}{|\frac{2\pi}{T}m|^{1+n}}, \tag{3.14}$$

that is,

$$|\epsilon| \leq AT^{1+n}\zeta(1+n). \tag{3.15}$$

The function ζ is the Riemann zeta function and A is a constant. In this case the approximation error decays like $\mathcal{O}(T^{-1-n})$.

From Figure 2.6, we can ascertain that the cubic spline model that we chose corresponds to an almost bandlimited function since it decays like $\mathcal{O}(|\omega|^4)$.

3.4.3 Artificial edges

In practice, the rays through the volume have finite lengths because of the finite size of the volume. This means that the limits of the integral in the projection (3.4) are finite and determined by the intersections of the rays with the volume facets.

The consequence of the finite, different ray lengths are artificial edges and artificial gradients in the projection image. They can be avoided in two ways. One is the normalization of the projection by the in-volume ray length, while the other is the projections of the volume after having removed its mean value. We follow the latter approach.

3.5 CST approach

We now describe the method that computes the FT of the projection by a CST-based extraction of the corresponding central slice of the 3D FT of the volume.

Let us write the 2D FT of the projection $p_{\boldsymbol{\mu}}$ from (3.4)

$$\hat{p}_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} f\left(R^{-1}(\Lambda^{-1}\mathbf{k} - \mathbf{t}) + \xi\, R^{-1}\Lambda^{-1}\mathbf{e}_3\right) \mathrm{d}\xi \right) e^{-j\langle \boldsymbol{\omega}, \mathbf{k}\rangle}\, \mathrm{d}^2 k, \quad (3.16)$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2, 0) \in \mathbb{R}^2 \times \{0\}$ is a spatial-frequency vector and where $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^\top \mathbf{y}$ stands for the inner product between the vectors $\mathbf{x}$ and $\mathbf{y}$.

We rewrite (3.16) as

$$\hat{p}_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = \int_{\mathbb{R}^3} f\left(R^{-1}\left(\Lambda^{-1}\mathbf{r} - \mathbf{t}\right)\right) e^{-j\langle \boldsymbol{\omega}, \mathbf{r}\rangle}\, \mathrm{d}^3 r, \quad (3.17)$$

where $\mathbf{r} = \mathbf{k} + \xi\, \mathbf{e}_3$.

Let us change the integration vector $\mathbf{r}$ by $\mathbf{s} = (s_1, s_2, s_3)$ as

$$\mathbf{s} = R^{-1}(\Lambda^{-1}\mathbf{r} - \mathbf{t}), \quad \mathrm{d}^3 s = |\det \Lambda^{-1}| \mathrm{d}^3 r. \quad (3.18)$$

Plugging $\mathbf{r}$ and $\mathrm{d}\mathbf{r}$ from (3.18) into (3.17) yields

$$\hat{p}_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = |\det \Lambda| \int_{\mathbb{R}^3} f(\mathbf{s})\, e^{-j\langle \boldsymbol{\omega}, \Lambda(R\mathbf{s} + \mathbf{t})\rangle}\, \mathrm{d}^3 s, \quad (3.19)$$

which we rewrite as

$$\hat{p}_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = |\det \Lambda| \left(\int_{\mathbb{R}^3} f(\mathbf{s})\, e^{-j\langle (\Lambda R)^\top \boldsymbol{\omega}, \mathbf{s}\rangle}\, \mathrm{d}^3 s \right) e^{-j\langle \boldsymbol{\omega}, \Lambda \mathbf{t}\rangle}, \quad (3.20)$$

where the term in parenthesis represents the 3D FT $\hat{f}$ of the volume f evaluated at the 3D frequency $(\Lambda R)^\top \boldsymbol{\omega}$. The 2D FT of the projection $p_{\boldsymbol{\mu}}$ is therefore given by

$$\hat{p}_{\boldsymbol{\mu}}(\boldsymbol{\omega}) = |\det \Lambda| \, \hat{f}\left((\Lambda R)^\top \boldsymbol{\omega}\right) e^{-j\langle \boldsymbol{\omega}, \Lambda \mathbf{t} \rangle}. \tag{3.21}$$

Equation (3.21) is a matrix form of the CST suited to our parameterization. It tells us how we can obtain the 2D FT of a projection at pose $\boldsymbol{\mu}$ using the 3D FT of the volume. It comprises the extraction of the central slice at pose determined by R and Λ from the 3D FT, the slice phase shift determined by Λ and $\mathbf{t}$, and the slice scaling by $|\det \Lambda|$. Since we have discrete data at our disposal, we replace the FT by the discrete FT (DFT). The slice extraction involves the interpolation of the real and imaginary parts of the 3D DFT of the volume. We model each part of the complex 3D DFT by a B-spline 3D data model (2.24).

We make use of the CST approach although it requires data interpolation in the FT domain. To achieve a more accurate interpolation in the FT domain, we suppose that the data to deal with are nearly space-limited and we choose the basis functions for the interpolation such that the aliasing in the space domain is avoided, as explained in Section 2.1.1.

3.5.1 CST approach vs. ray casting: Speed

The space-based method of volume-to-image registration necessitates one projection of the volume per iteration. Equation (3.10) requires the volume interpolation all the way along the ray through the volume. The cost of computing a projection of size $N \times N$ pixel of the volume of size $N \times N \times N$ voxel is therefore $\mathcal{O}(N^3)$.

The frequency-based method demands the computation of the 3D DFT of the volume and of the 2D DFTs of the reference projections. They are computed only once as a part of data preprocessing. The computational cost of a 3D Fast Fourier Transform (FFT) is $\mathcal{O}(N^3 \log N)$, while the cost of a 2D FFT computation is $\mathcal{O}(N^2 \log N)$. This approach also demands the extraction of one central slice of the 3D DFT per iteration. Equation (3.21) requires the interpolation to be performed at only one point of the real part of $\hat{f}$ and at only one point of the imaginary part of $\hat{f}$, per image-plane pixel. The computational cost of the slice extraction is therefore $\mathcal{O}(N^2)$.

The CST-based simulated 2D data for the dissimilarity evaluation are therefore less costly than those obtained by ray casting.

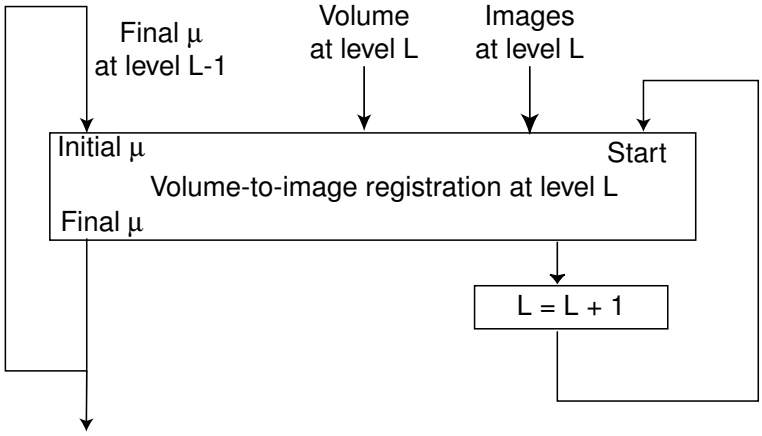


Figure 3.4: Space-domain multiresolution strategy.

3.6 Multiresolution strategy

The coarse-to-fine data-processing approach is a standard tool for improving the robustness of optimization algorithms. The data at coarser resolutions are smoothed versions of the data at full resolution. The registration using data with no or just a few details augments the chances of reaching the global optimum instead of getting trapped on the way into a local one. The strategies that are based on reducing the data size speed up the optimization. Different multiresolution strategies can be developed for both the space- and frequency-based registration. Here, we propose two that give a good tradeoff between accuracy and time.

3.6.1 Space domain

Figure 3.4 shows a block-diagram of our space-domain multiresolution strategy. Given a volume and a set of reference images, we propose to precompute their spline L_2 pyramids (Section 2.3). We register the volume to the images starting from their coarsest resolution. We use the final estimate obtained at some pyramid level to resume the registration at the next finer one. We repeat the registration procedure at finer resolutions until we reach a resolution that provides a good tradeoff between accuracy and time [47].

The cost of an iteration at some coarse data resolution is negligible since we

have few data to process, and it increases as we go towards finer resolutions. Consequently, we can afford performing many iterations at coarse resolutions but only few of them at finer ones. This means that the total time required to achieve the registration in a multiresolution fashion is much shorter than if we performed the registration at full resolution.

3.6.2 Frequency domain

Thanks to Parseval's theorem, the least-squares criterion can also be written in the Fourier domain.

Let us rewrite the cost function from (3.3) as

$$S(\boldsymbol{\mu}) = \frac{1}{2} \sum_{q=1}^{Q} \sum_{i=1}^{M} \hat{r}_{q,i}^{*}\, \hat{r}_{q,i}, \tag{3.22}$$

where $\hat{r}_{q,i} = [\hat{\mathbf{r}}_q(\boldsymbol{\mu})]_i$ and where $\hat{r}^*$ stands for the complex conjugate of $\hat{r}$. To perform the registration in the frequency domain using a multiresolution processing, we propose to weight the cost function by the intensity assigned to each pixel of a 2D weighting function w as follows:

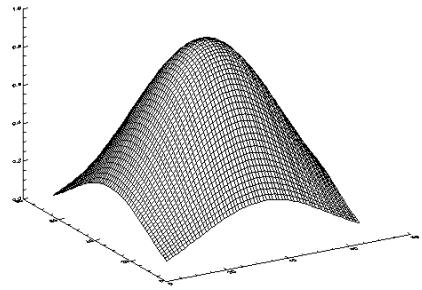
$$S(\boldsymbol{\mu}) = \frac{1}{2} \sum_{q=1}^{Q} \sum_{i=1}^{M} w_i\, \hat{r}_{q,i}^{*}\, \hat{r}_{q,i}, \tag{3.23}$$

where w_i are the components of a vector-form w.

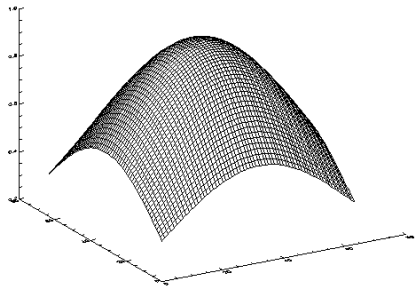
We employ a 2D Gaussian-shape weighting function centered at the origin, with a tunable standard deviation. This weighting function acts as a low-pass filter since it rewards lower data frequencies and penalizes higher ones. In this case, a coarse-to-fine processing can be achieved by increasing the standard deviation of the Gaussian, which means, by increasing the contribution of higher frequencies to the cost function. The weighting function for three different values of the standard deviation of the Gaussian is shown in Figure 3.5. In this work, we have often used a Gaussian-shape weighting function of size $P \times P$ pixels with the standard deviation $P/2$.

As we have seen, the registration speed is related to the size of the processed data. Contrary to the space-domain multiresolution strategy, the frequency-domain one does not increase the registration speed since it does not reduce the data size. A reduction could be achieved only by computing a partial sum in

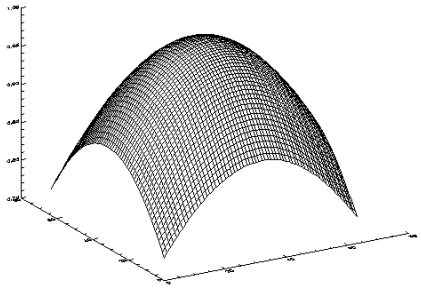
(3.23), but this would result in a hard-edge filter which would introduce ringing-associated artifacts, perhaps to the point of creating unwanted additional local minima.



(a) The standard deviation equal to $P/3$.



(b) The standard deviation equal to $P/2$.



(c) The standard deviation equal to P.

Figure 3.5: 2D Gaussian as the weighting function for the frequency-domain multiresolution strategy. Three different values for the standard deviation of the Gaussian of size $P \times P$ pixels ($P = 64$).

3.7 Iterative optimization

We can classify existing iterative methods for solving an optimization problem into two classes. One set of optimizers requires the gradient of the cost function to be available, either analytically or via a finite difference. The cost function has to be smooth, and the gradient-based optimizers can fail if the cost function has discontinuities or irregularities. Another set of optimizers does not require the gradient. They are used if the gradient is not available or if the cost function is not smooth enough. The study of optimizers is research-in-progress. As soon as the cost function is complicated enough, no optimizer is capable of finding the global optimum, even if some succeed more often than others [49].

We propose to implement a gradient-based optimizer along with a continuous B-spline volume model (Section 2.2), which makes the cost function (3.3) smooth and makes the exact computation of the gradients possible.

3.7.1 Gradient-based algorithms

Newton method

This algorithm [8, 49] updates the estimation of the parameters $\boldsymbol{\mu}$ for the next, $(k+1)$-th iteration, by minimizing the local quadratic model of the cost function S around the current estimation $\boldsymbol{\mu}^{(k)}$

$$J(\boldsymbol{\mu}) = S(\boldsymbol{\mu}^{(k)}) + \boldsymbol{\nabla} S(\boldsymbol{\mu}^{(k)})^{\top}(\boldsymbol{\mu} - \boldsymbol{\mu}^{(k)}) + \frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{\mu}^{(k)})^{\top}\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}^{(k)})(\boldsymbol{\mu} - \boldsymbol{\mu}^{(k)}). \tag{3.24}$$

If $\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}^{(k)})$ is positive definite, the minimizer exists and is the unique solution of $\boldsymbol{\nabla} J(\boldsymbol{\mu}) = \mathbf{0}$

$$\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} - (\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}^{(k)}))^{-1}\, \boldsymbol{\nabla} S(\boldsymbol{\mu}^{(k)}). \tag{3.25}$$

However, if $\boldsymbol{\mu}^{(k)}$ is far from the solution, $\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}^{(k)})$ could have negative eigenvalues, in which case, the model given by (3.24) does not have local minimizers, and the algorithm may converge to local maxima or saddle points of (3.24). There are two potential reasons for this failure: either the Newton direction fails to be a direction of decrease of S, or the step length is too large. In return, the Newton method is extremely efficient when near to a minimizer; it converges quadratically towards it.

Method of steepest descent

This method [8, 49] updates the current-iteration estimate as follows:

$$\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} - \Gamma \, \boldsymbol{\nabla} S(\boldsymbol{\mu}^{(k)}), \tag{3.26}$$

where Γ is a scaling diagonal matrix. Unlike the Newton method, the method of steepest descent is globally convergent. This means that it does not require the initial estimate to be close to a local minimizer. However, it does not guarantee the next estimate to be nearer to a solution than the current one, even if the latter is very close to a solution. This method may be slow since it converges linearly. Its success depends on the choice of Γ. One appropriate choice for the diagonal elements of Γ, $\gamma_i, i = 1, 2, \ldots, N$, is $\gamma_i = \eta^m$, where $\eta \in (0, 1)$ and $m \geq 0$ is the smallest nonnegative integer such that there is a sufficient decrease in S.

Newton-like method

In this section, we describe a method that achieves the best tradeoff between the robust but generally inefficient method of steepest descent and the efficient but not robust Newton method [8, 49].

Let H be a modified Hessian such that the diagonal elements of the true Hessian are multiplied by a factor λ while its off-diagonal elements are not changed

$$[H(\boldsymbol{\mu})]_{i,j} = [\boldsymbol{\nabla}^2 S(\boldsymbol{\mu})]_{i,j} \, (1 + \lambda \, \delta_{i,j}), \tag{3.27}$$

where $\delta_{i,j}$ is the Kronecker symbol and $i, j \in \{1, 2, \ldots, N\}$.

Then, the Newton-like optimization algorithm can be described by

$$\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} - (H(\boldsymbol{\mu}^{(k)}))^{-1} \, \boldsymbol{\nabla} S(\boldsymbol{\mu}^{(k)}). \tag{3.28}$$

Equation (3.28) approximates (3.26) for $\lambda \to +\infty$. Similarly, it approximates (3.25) for $\lambda \to 0$. The parameter λ is tuned to provide a smooth transition from the steepest-descent algorithm (used in the beginning) to the Newton algorithm (used when approaching to the solution).

Levenberg-Marquardt method

The LM optimization algorithm is defined in [76] as the Newton-like method that is used with a least-squares cost function; in this case, a specific approximation of

the Hessian leads to even better convergence properties. For efficiency reasons, we apply the LM algorithm to optimize the least-squares cost function given by (3.3). We implement the following expressions for the gradient and the Hessian:

Gradient. The gradient of the least-squares cost function from (3.3) is

$$\boldsymbol{\nabla} S(\boldsymbol{\mu}) = \sum_{q=1}^{Q} (R_q^{'}(\boldsymbol{\mu}))^{\top} \mathbf{r}_q(\boldsymbol{\mu}), \tag{3.29}$$

where $R_q^{'}$ is the $M \times N$ Jacobian matrix of $\mathbf{r}_q$

$$[R_q^{'}(\boldsymbol{\mu})]_{i,j} = \frac{\partial r_{q,i}}{\partial \mu_j},\ 1 \leq i \leq M,\ 1 \leq j \leq N. \tag{3.30}$$

To determine $R_q^{'}$, we compute the first derivative of the spline data model given by (2.25) for space-domain interpolation, and by (2.26) for frequency-domain interpolation.

Hessian. The Hessian of the least-squares cost function from (3.3) is given by

$$\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}) = \sum_{q=1}^{Q} (R_q^{'}(\boldsymbol{\mu}))^{\top} R_q^{'}(\boldsymbol{\mu}) + \sum_{q=1}^{Q} (R_q^{''}(\boldsymbol{\mu}))^{\top} \mathbf{r}_q(\boldsymbol{\mu}). \tag{3.31}$$

The second-order term in (3.31) is too costly to be computed in practice. Besides, the rightmost sum is negligible for small residuals (because the value of the cost function at a minimizer is small). Also, the influence of the second derivative terms can be destabilizing in some cases. We thus employ the following approximation of the elements of the Hessian matrix which ignores them [8, 49, 76]:

$$\boldsymbol{\nabla}^2 S(\boldsymbol{\mu}) \approx \sum_{q=1}^{Q} (R_q^{'}(\boldsymbol{\mu}))^{\top} R_q^{'}(\boldsymbol{\mu}). \tag{3.32}$$

This is known as the Gauss-Newton approximation [8, 49].

Chapter 4

REGISTRATION OF A 3D CT TO C-ARM IMAGES

In this chapter, we present an application of our space-based algorithm in CAOS. This is the registration of a CT volume to a set of C-arm images of a patient for improved 3D CT-based navigation.

A short overview of navigation methods for CAOS is given in Section 4.1. In the same section, we discuss the thesis contribution to an improved navigation.

In Section 4.2, we adapt the space-based algorithm described in Chapter 3 to the cone-beam C-arm projection geometry. We also give the details left out from Chapter 3.

The real CT/C-arm data used for validating of the algorithm are described in Section 4.3. The validation which is performed in two ways is shown in Section 4.4. In both cases, we use volume/image data whose true alignment is known *a priori*. This ensures that our validation is objective. First, we test the performance of the algorithm when registering the CT volume to synthetic, simulated C-arm images, computed for a known CT pose. This fully controlled environment allows for a truly objective validation of the quality of our algorithm. Second, we test the performance of the algorithm when registering the CT volume to a set of real C-arm images based on fiducial markers. As the location of these markers is known only within some error margin, the validation is less objective.

Figure 4.1: GE LightSpeed Ultra CT scanner.

4.1 Computer-assisted orthopedic surgery

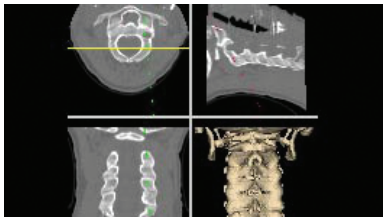
Currently available systems for computer-assisted surgery include robotic and volume/image-guided procedures. These systems are used in various surgical fields. Here, we focus on orthopedic surgery [2, 7, 24, 94]. Surgical navigation means a visualization of the position of surgical tools and implants relative to the patient (e.g., spine, pelvis) during the intervention. Standard orthopedic-surgery systems are based on preoperative CT volume-based navigation; they have been used successfully for spinal pedicle screw insertion, total hip arthroplasty, total knee arthroplasty, pelvic osteotomy, and reconstruction of knee cruciate ligaments [7, 94]. Intraoperative C-arm image-based navigation is a recently developed technique that has shown promising results in early clinical trials to perform spinal pedicle screw insertion, distal locking of femoral nails, and femoral fracture reduction [7, 69]. There have been many attempts to combine good properties of the two methods to achieve an improved 3D CT-based navigation [2, 34, 36, 51, 53, 74, 81, 100, 117]. One of the contributions of this thesis is the development of a method for improved 3D CT-based navigation thanks to a set of intraoperative C-arm images.

4.1.1 3D CT-based navigation

This method offers a display of surgical tools and implants in the 3D CT that is acquired preoperatively, using a CT scanner (Figure 4.1), and used by the surgeon for planning the intervention. CT-based navigation requires a registration of the CT volume to the patient. This registration is based either on matching points (fiducials implanted onto the subject before acquiring the data, or anatomical landmarks manually identified on the CT and on the subject), or on

matching surfaces (matching of the surface extracted from the CT to the surface data points collected intraoperatively using touch or ultrasonic probes). A very good characteristic of this method is a possibility of tracking the tools in a 3D space that provides information about the structure of the object (Figure 4.2).

(a) 3D CT-guided spinal surgery.



(b) Typical display: the transversal (upper left corner), sagittal (upper right corner), and frontal (lower left corner) slices of the spine CT volume.

Figure 4.2: 3D CT-based navigation that provides information about the actual position of the tools in the transversal, sagittal, and frontal planes, as well as about any deviation from the original plan (from http://www.medivision.ch/).

Drawbacks of this navigation method are found in the registration part. For example, anatomical-landmark-based registration is not reproducible, fiducial-based registration is invasive, and the accuracy of surface-based registration is affected by the accuracy of the surface extraction and by the quantity and quality of the intraoperative data collection. We avoid these drawbacks and proceed instead with the methods of Section 4.1.3.

Figure 4.3: Siemens ISO-C C-arm.

4.1.2 C-arm image-based navigation

With the type of navigation described in the present section, tools are displayed in intraoperative images of the patient obtained by a mobile C-arm device (Figure 4.3) at different viewing angles. These images are displayed on a computer screen. Computed projections of the tools are displayed in each image (Figure 4.4). An update rate of 10 Hz enables real-time navigation in up to four C-arm images simultaneously in [43]. These images are registered to the patient by tracking the patient and the C-arm image intensifier simultaneously [43, 69] with a position sensor (Figure 4.5).

A setup for C-arm-image-based navigation has been proposed in [43, 69]; it is the one used in this thesis, and is shown in the upper part of Figure 4.6. The tool, the patient, and the C-arm image intensifier are equipped with light-emitting diode (LED) markers that define the local COordinate Systems (COS). The position of the LED markers is tracked by an optoelectronic camera that defines the Reference COS (R-COS). The position of the the tip of the tool in the 2D, pixel-based C-arm Image COordinate System (I-COS) is computed through a chain of coordinate transformations between the local COS. The coordinate of the tip of the tool $\mathbf{v}_{\mathrm{T}}$ is first transformed from the Tool COS (T-COS) to the Patient COS (P-COS). The resulting coordinate is then transformed from the P-COS to the C-arm image intensifier COS (A-COS). This coordinate is finally projected from the 3D A-COS to the 2D I-COS using a cone-beam projection

model.

The advantage of this type of navigation is the availability of C-arm machines. Thanks to its mobility and easy manipulation, C-arm has become a standard piece of equipment for orthopedic surgery. However, to compensate the lack in three-dimensionality (missing C-arm image depth), images at many different C-arm orientations should be displayed to the surgeon. This is a drawback of the method since it involves an augmented radiation dosage of the surgical staff and of the patient. This thesis contributes to overcoming this drawback, as described in the next section.

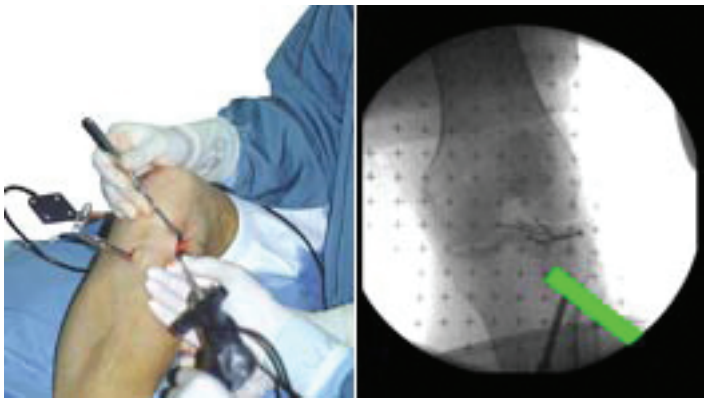


Figure 4.4: The C-arm image-based navigation provides information about the actual position of the tools in any anatomical region (from http://www.medivision.ch/). Left: Knee surgery. Right: Tool projection in a C-arm image of the knee.

Figure 4.5: Northern Digital Optotrak 3020.

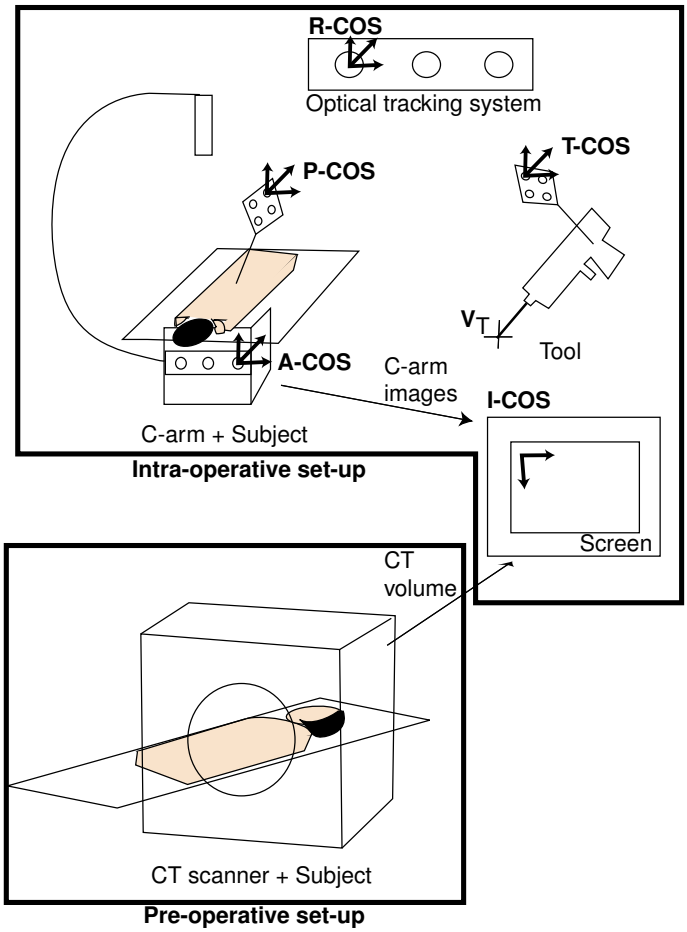


Figure 4.6: Set-up for an improved 3D CT-based navigation.

4.1.3 Improved 3D CT-based navigation

The idea is to develop a strategy for CT volume-based navigation where the preoperative 3D CT is registered to as few as possible intraoperative C-arm images of the patient. This registration means a computation of the R-COS pose of the CT volume with respect to the C-arm images. It makes possible the display of tools in the CT volume (Figure 4.6).

Novelties that our algorithm brings. Our algorithm for registering a CT to C-arm images is projection-based, and therefore, is similar to [74, 75, 116, 117]. Contrary to these previously published methods, ours takes advantage of a continuous image/volume model based on cubic B-splines for computing the following three elements: the projections of the volume, the gradient of the similarity measure, and the multiresolution data pyramids. In addition, we compute fast projections in a one-step procedure that does not complicate the optimization of the similarity measure. In return, the optimization is performed simultaneously in all unknown parameters.

Given at least two poses of the C-arm image plane in the R-COS, our algorithm computes the pose of the CT volume in the R-COS at which the volume is in registration with the two C-arm images. In contrast, the C-arm image pose was unknown (no tracking system) and only one C-arm image was provided in [74, 75, 116, 117]. The performance of our algorithm is therefore potentially better than that of [74, 75, 116, 117].

4.2 Registration algorithm

We register a CT volume to a set of C-arm images of a patient using the space-based algorithm shown in Figure 4.7. The input data for the algorithm are: 1) the CT volume, 2) the C-arm images, 3) the geometrical data of the cone-beam C-arm projection model, and 4) the initial volume pose.

4.2.1 Volume pose

In Chapter 3, we defined the volume pose in the customary way of 3D EM. In this chapter, we define the volume pose as described in [98].

Let $\mathbf{n} = (n_1, n_2, n_3, 1)$ be the homogenous 3D index of a point in the volume, let $\mathbf{v} = (v_1, v_2, v_3, 1)$ be the coordinate of this point in the R-COS, and let $B_{\boldsymbol{\mu}}$ be the transformation from $\mathbf{n}$ to $\mathbf{v}$ that is determined by the parameters $\boldsymbol{\mu}$ of the volume pose in the R-COS, that is, $\mathbf{v} = B_{\boldsymbol{\mu}}\mathbf{n}$. We express the volume pose by the three Euler rotation angles φ, θ, and ψ, that are to be performed in the same way as in [98], and by the three translations Δx, Δy, and Δz. We write $\boldsymbol{\mu} = (\varphi, \theta, \psi, \Delta x, \Delta y, \Delta z)$. Given the parameters $\boldsymbol{\mu}$, the voxel size $\lambda_x \times \lambda_y \times \lambda_z$, and the 3D index of the volume origin $\mathbf{c}$, we propose to write the transformation $B_{\boldsymbol{\mu}}$ as follows:

$$B_{\boldsymbol{\mu}} = T\Lambda C^{-1} R_x R_y R_z C. \tag{4.1}$$

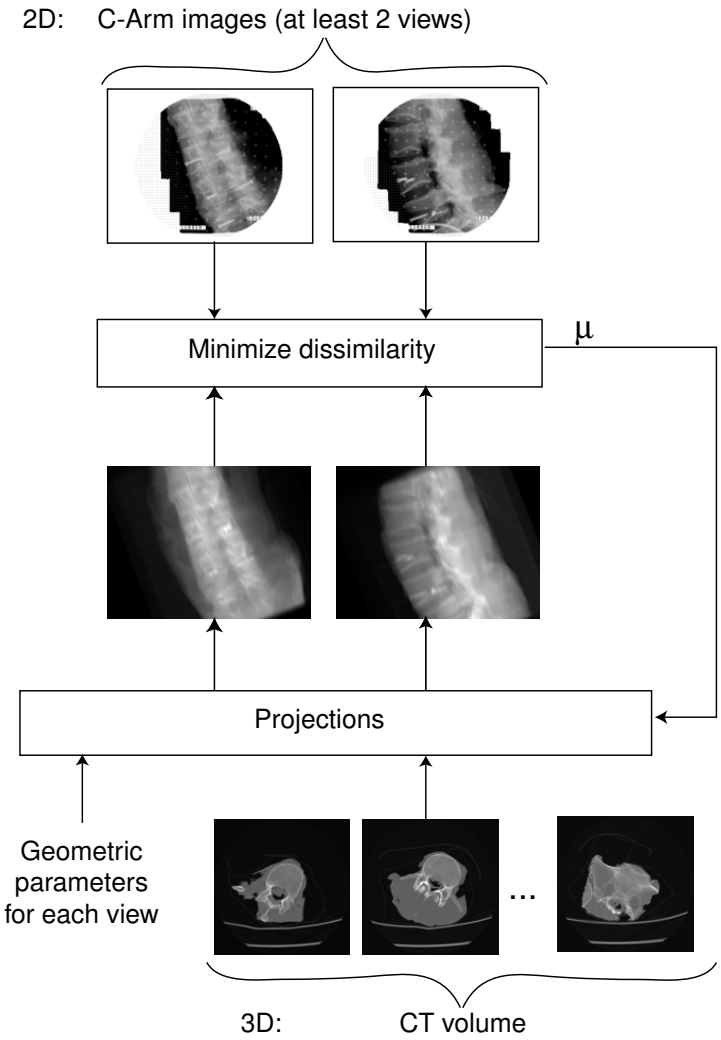


Figure 4.7: Space-based algorithm for registering a 3D CT to C-arm images.

The matrix

$$C = \begin{pmatrix} 1 & 0 & 0 & -[\mathbf{c}]_1 \\ 0 & 1 & 0 & -[\mathbf{c}]_2 \\ 0 & 0 & 1 & -[\mathbf{c}]_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.2}$$

makes the origin of the volume be the center of rotation. The matrix

$$R_z = \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 & 0 \\ \sin(\psi) & \cos(\psi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.3}$$

rotates the volume by the angle ψ around its z-axis. Similarly,

$$R_y = \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.4}$$

rotates the volume by the angle θ around its new y-axis, while

$$R_x = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) & 0 \\ 0 & \sin(\varphi) & \cos(\varphi) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.5}$$

rotates the volume by the angle φ around its new x-axis. C^{-1} shifts back the index of the volume origin, and

$$\Lambda = \begin{pmatrix} \lambda_x & 0 & 0 & 0 \\ 0 & \lambda_y & 0 & 0 \\ 0 & 0 & \lambda_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.6}$$

scales the volume indexes by the voxel size. Finally,

$$T = \begin{pmatrix} 1 & 0 & 0 & \Delta x \\ 0 & 1 & 0 & \Delta y \\ 0 & 0 & 1 & \Delta z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.7}$$

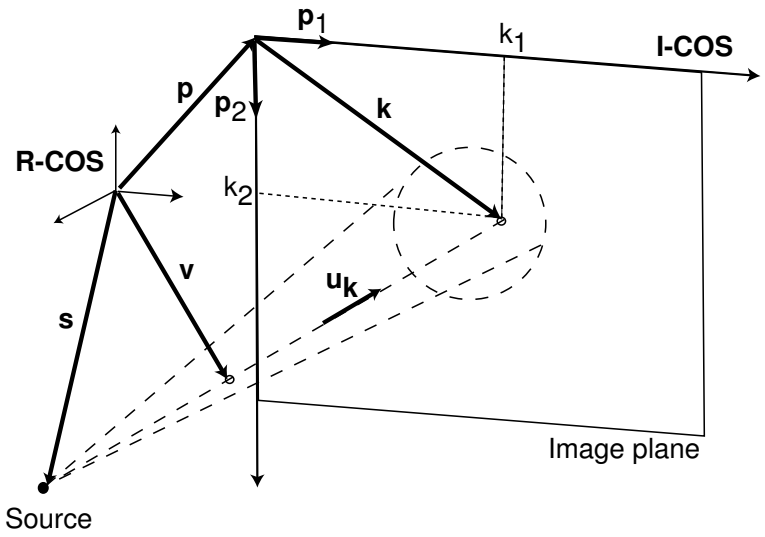


Figure 4.8: Cone-beam projection geometry.

translates the volume by Δx, Δy, and Δz along its new x-,y-, and z-axis, respectively.

The sense of rotation for φ, θ, and ψ, has been chosen so as to correspond to the conventions of a right-handed coordinate system.

4.2.2 Cone-beam projection geometry

We define the pose of the image plane (Figure 4.8) using the R-COS coordinates of a point $\mathbf{p}$ and the two orthonormal unit vectors $\mathbf{p}_1$,$\mathbf{p}_2$ inside the plane ($\mathbf{p}_1 \perp \mathbf{p}_2$, and $\|\mathbf{p}_1\| = \|\mathbf{p}_2\| = 1$). Let $\mathbf{s}$ be the coordinate of the illumination source in the R-COS and let $p_x \times p_y$ be the pixel size.

We propose to use a perspective projection model. Let the projection of a point, with the coordinate $\mathbf{v}$ in the R-COS, onto the image plane, be determined in the I-COS by a 2D index $\mathbf{k} = (k_1, k_2, 0, 1)$. Given $\mathbf{v}$, we can find $\mathbf{k}$ by solving the non-linear equation

$$\frac{A\mathbf{k} - \mathbf{s}}{\|A\mathbf{k} - \mathbf{s}\|} = \frac{\mathbf{v} - \mathbf{s}}{\|\mathbf{v} - \mathbf{s}\|}, \tag{4.8}$$

where

$$A = A_2 A_1, \quad A_1 = \begin{pmatrix} p_x & 0 & 0 & 0 \\ 0 & p_y & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} [\mathbf{p}_1]_1 & [\mathbf{p}_2]_1 & 0 & [\mathbf{p}]_1 \\ [\mathbf{p}_1]_2 & [\mathbf{p}_2]_2 & 0 & [\mathbf{p}]_2 \\ [\mathbf{p}_1]_3 & [\mathbf{p}_2]_3 & 1 & [\mathbf{p}]_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.9}$$

Note that the two sides of (4.8) are two different ways to express the unit vector on the ray from the source towards a point in the image plane with the I-COS index $\mathbf{k}$. Let us denote this unit vector by $\mathbf{u_k}$.

4.2.3 Projection

Let us suppose that the volume is sitting between the source and the image plane. The R-COS coordinate $\mathbf{v}$ of any point of the volume on the ray, which is determined by the unit vector $\mathbf{u_k}$, can be expressed as

$$\mathbf{v} = \mathbf{s} + \xi\ \mathbf{u_k}, \ \ \xi \in [a_{\mathbf{k};\boldsymbol{\mu}}, b_{\mathbf{k};\boldsymbol{\mu}}] \subset [0, \|A\mathbf{k} - \mathbf{s}\|], \tag{4.10}$$

where $a_{\mathbf{k};\boldsymbol{\mu}}$ and $b_{\mathbf{k};\boldsymbol{\mu}}$ stand for the value of ξ at the intersections of the ray with the volume facets. This coordinate $\mathbf{v}$ corresponds to the 3D index in the volume $\mathbf{n} = B_{\boldsymbol{\mu}}^{-1}\mathbf{v}$.

The cone-beam projection of the volume f (mean-value removed, as explained in Section 3.4.3) is given by

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \int_{a_{\mathbf{k};\boldsymbol{\mu}}}^{b_{\mathbf{k};\boldsymbol{\mu}}} f\left(B_{\boldsymbol{\mu}}^{-1}\ \left(\mathbf{s} + \xi\ \frac{A\mathbf{k} - \mathbf{s}}{\|A\mathbf{k} - \mathbf{s}\|}\right)\right)\ \mathrm{d}\xi. \tag{4.11}$$

We write the projection (4.11) in the following simplified form

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \int_a^b f(\mathbf{n}_0 + \xi\,\mathbf{n})\,\mathrm{d}\xi. \tag{4.12}$$

This form is identical to (3.5); the difference between the parallel-beam projection and the present cone-beam projection lies in the value of the variables given below. In particular, note that $\mathbf{n}$ is nonlinear with respect to $\mathbf{k}$.

$$\mathbf{n}_0 = B_{\boldsymbol{\mu}}^{-1}\mathbf{s}, \quad \mathbf{n} = B_{\boldsymbol{\mu}}^{-1}\tfrac{A\mathbf{k}-\mathbf{s}}{\|A\mathbf{k}-\mathbf{s}\|}, \quad a = a_{\mathbf{k};\boldsymbol{\mu}}, \quad b = b_{\mathbf{k};\boldsymbol{\mu}}. \tag{4.13}$$

4.2.4 Fast projection

We derive now a fast projection method for the cone-beam projection. This derivation follows the same path as was indicated in Section 3.4.1 in the context of a parallel-beam projection. First, we change the integration variable ξ in (4.12) by the following three variables $\xi_j, j \in \{1,2,3\}$

$$\xi_j = [\mathbf{n}_0]_j + \xi\,[\mathbf{n}]_j \longrightarrow \mathrm{d}\xi_j = [\mathbf{n}]_j\,\mathrm{d}\xi. \tag{4.14}$$

After expressing ξ and $\mathrm{d}\xi$ from (4.14), we replace them in (4.12). After some algebraic manipulations on (4.12), we obtain three expressions for the projection

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \lambda_j \int_{a_j}^{b_j} f(\mathbf{n}_0^j + \xi_j\,\mathbf{n}^j)\,\mathrm{d}\xi_j,\ j \in \{1,2,3\}, \tag{4.15}$$

where $\lambda_j = \frac{1}{[\mathbf{n}]_j}$, and

$$\mathbf{n}_0^j = \mathbf{n}_0 - \tfrac{[\mathbf{n}_0]_j}{[\mathbf{n}]_j}\mathbf{n},\quad \mathbf{n}^j = \tfrac{1}{[\mathbf{n}]_j}\mathbf{n},\quad a_j = [\mathbf{n}_0]_j + a\,[\mathbf{n}]_j,\quad b_j = [\mathbf{n}_0]_j + b\,[\mathbf{n}]_j. \tag{4.16}$$

As mentioned in Section 3.4.1, we propose to chose that j for which $|\lambda_j|$ is minimum. Since λ_j depends on $\mathbf{k}$, we have to chose j for each image pixel, independently.

We adapt the phase as described in Section 3.4.1. We can choose any finite $\Delta\xi_j$ and still write (4.15) as

$$p_{\boldsymbol{\mu}}(\mathbf{k}) = \lambda_j \int_{a_j-\Delta\xi_j}^{b_j-\Delta\xi_j} f\left(\mathbf{n}_0^j + (\xi_j + \Delta\xi_j)\,\mathbf{n}^j\right)\,\mathrm{d}\xi_j,\ j \in \{1,2,3\}. \tag{4.17}$$

This leads to the following discrete version:

$$p_{\boldsymbol{\mu}}(\mathbf{k}) \approx \lambda_j \sum_{m \in \mathbb{D}_j} f(\mathbf{n}_0^j + (m + \Delta\xi_j)\,\mathbf{n}^j),\ j \in \{1,2,3\}, \tag{4.18}$$

where $\mathbb{D}_j \subset \mathbb{Z}$, and where $\Delta\xi_j$ is chosen such that $\left[\mathbf{n}_0^j + \Delta\xi_j\,\mathbf{n}^j\right]_j = 0$. We note again that the evaluation of f requires the interpolation of the volume at a point with one integer and only two real coordinates. This means that we have replaced the required 3D interpolation of the volume in (4.12) by its faster 2D interpolation. An example of applying the shearing method for computing a 1D cone-beam projection of an image is shown in Figure 4.9.

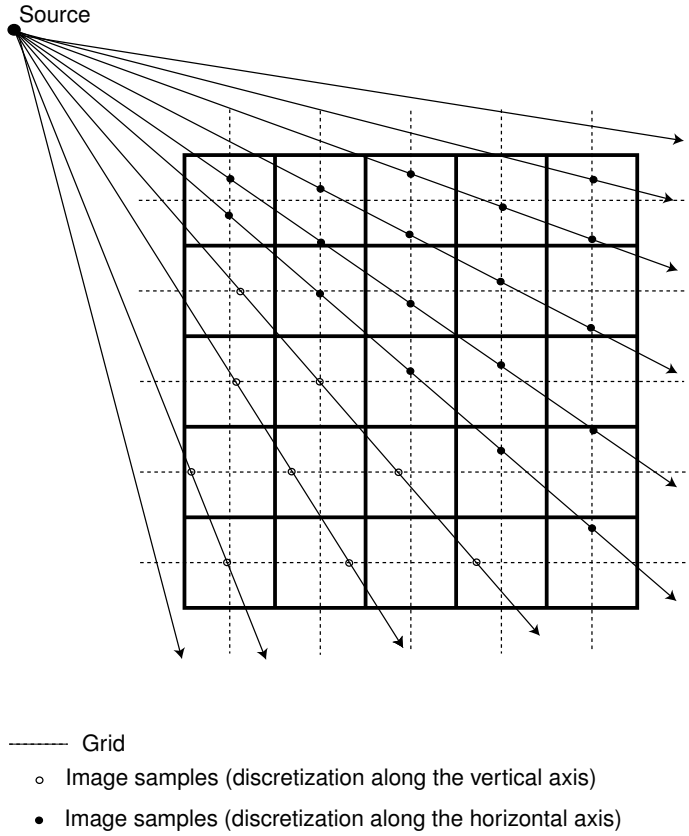


Figure 4.9: Shearing approach to 1D cone-beam projection of an image.

We observe two image regions in this figure. In one, we sample the image along the horizontal axis, while in the other, the sampling is performed along the vertical axis. This means that in the 3D-to-2D case we can find image pixels with the coordinate $\mathbf{k}$ for which j is same. This can speed up the implementation of the projection equation.

4.2.5 Cost function

We chose to minimize the following measure of dissimilarity between the C-arm images and the projections of the volume:

$$S(\boldsymbol{\mu}) = \frac{1}{2} \sum_{q=1}^{Q} \frac{1}{\mathrm{card}(\mathbb{D}_q)} \sum_{\mathbf{k} \in \mathbb{D}_q} \left(\frac{p_{\boldsymbol{\mu};q}(\mathbf{k}) - P_{\boldsymbol{\mu};q}}{\sigma_{\boldsymbol{\mu};q}} - \frac{p_q(\mathbf{k}) - P_q}{\sigma_q} \right)^2, \qquad (4.19)$$

where p_q is the q-th C-arm image, $p_{\boldsymbol{\mu};q}$ is the q-th projection of the volume at pose $\boldsymbol{\mu}$, $\mathbb{D}_q$ is an arbitrary mask for the q-th image, and Q is the number of C-arm viewing angles.

This measure is robust to differences in the intensity range of the C-arm images and the corresponding projections. The robustness is achieved by normalizing mean-subtracted images by their standard deviations. The mean and the standard deviation on the domain $\mathbb{D}_q$ of the C-arm image are denoted by P_q and σ_q, respectively, ($P_q = \frac{1}{\mathrm{card}(\mathbb{D}_q)} \sum_{\mathbf{k} \in \mathbb{D}_q} p_q(\mathbf{k})$, $\sigma_q^2 = \frac{1}{\mathrm{card}(\mathbb{D}_q)} \sum_{\mathbf{k} \in \mathbb{D}_q} (p_q(\mathbf{k}) - P_q)^2$). Similarly, $P_{\boldsymbol{\mu};q}$ and $\sigma_{\boldsymbol{\mu};q}$ denote the mean and the standard deviation on the domain $\mathbb{D}_q$ for the projection, respectively.

4.2.6 Multiresolution strategy

We apply the space-domain multiresolution strategy based on L_2-cubic spline volume/image pyramids to make the algorithm fast and robust, as explained in Section 3.6.1 .

4.2.7 Fast first derivative of the projection

We write the first derivative of the projection given by (4.17) with respect to the parameters $\boldsymbol{\mu}$ as

$$\begin{aligned} \frac{\partial p_{\boldsymbol{\mu}}(\mathbf{k})}{\partial \boldsymbol{\mu}} \quad = \quad & \frac{\partial \lambda}{\partial \boldsymbol{\mu}} \int_a^b f(\mathbf{n}_0 + (\xi + \Delta\xi)\,\mathbf{n})\,\mathrm{d}\xi + \\ & \lambda \left(\frac{\partial b}{\partial \boldsymbol{\mu}} f(\mathbf{n}_0 + (b + \Delta\xi)\,\mathbf{n}) - \frac{\partial a}{\partial \boldsymbol{\mu}} f(\mathbf{n}_0 + (a + \Delta\xi)\,\mathbf{n}) \right) + \\ & \lambda \int_a^b (\boldsymbol{\nabla} f(\mathbf{n}_0 + (\xi + \Delta\xi)\,\mathbf{n}))^\top \left(\frac{\partial \mathbf{n}_0}{\partial \boldsymbol{\mu}} + (\xi + \Delta\xi) \frac{\partial \mathbf{n}}{\partial \boldsymbol{\mu}} \right) \mathrm{d}\xi. \end{aligned} \qquad (4.20)$$

Note that we have omitted all j-s to simplify the expression; also, $a = a_j - \Delta\xi_j$ and $b = b_j - \Delta\xi_j$.

Next, we discretize (4.20)

$$\begin{aligned}\frac{\partial p_{\boldsymbol{\mu}}(\mathbf{k})}{\partial \boldsymbol{\mu}} \quad \approx \quad & \frac{\partial \lambda}{\partial \boldsymbol{\mu}} \sum_{m \in \mathbb{D}} f(\mathbf{n}_0 + (m + \Delta\xi)\,\mathbf{n}) + \\ & \lambda \left(\frac{\partial b}{\partial \boldsymbol{\mu}} f(\mathbf{n}_0 + (b + \Delta\xi)\,\mathbf{n}) - \frac{\partial a}{\partial \boldsymbol{\mu}} f(\mathbf{n}_0 + (a + \Delta\xi)\,\mathbf{n}) \right) + \\ & \lambda \sum_{m \in \mathbb{D}} (\boldsymbol{\nabla} f(\mathbf{n}_0 + (m + \Delta\xi)\,\mathbf{n}))^{\top} \left(\frac{\partial \mathbf{n}_0}{\partial \boldsymbol{\mu}} + (m + \Delta\xi) \frac{\partial \mathbf{n}}{\partial \boldsymbol{\mu}} \right),\end{aligned} \tag{4.21}$$

where $\mathbb{D} \subset \mathbb{Z}$. We introduce an approximation here. To reduce the computation complexity, we assume that the change of the integral limits a and b due to the change of $\boldsymbol{\mu}$ is negligible between two iterations. We therefore ignore $\frac{\partial b}{\partial \boldsymbol{\mu}}$ and $\frac{\partial a}{\partial \boldsymbol{\mu}}$ and implement the following equation:

$$\begin{aligned}\frac{\partial p_{\boldsymbol{\mu}}(\mathbf{k})}{\partial \boldsymbol{\mu}} \quad \approx \quad & \frac{\partial \lambda}{\partial \boldsymbol{\mu}} \sum_{m \in \mathbb{D}} f(\mathbf{n}_0 + (m + \Delta\xi)\,\mathbf{n}) + \\ & \lambda \sum_{m \in \mathbb{D}} (\boldsymbol{\nabla} f(\mathbf{n}_0 + (m + \Delta\xi)\,\mathbf{n}))^{\top} \left(\frac{\partial \mathbf{n}_0}{\partial \boldsymbol{\mu}} + (m + \Delta\xi) \frac{\partial \mathbf{n}}{\partial \boldsymbol{\mu}} \right).\end{aligned} \tag{4.22}$$

4.2.8 Approximated gradient and Hessian

The full expressions for the gradient and the Hessian are too complicated to be used in practice, unless we assume that the derivatives of $P_{\boldsymbol{\mu};q}$ and $\sigma_{\boldsymbol{\mu};q}$ with respect to $\boldsymbol{\mu}$ are equal to zero. Also, to improve the stability of the convergence of the optimizer, we ignore the second-order term in the Hessian as explained in Section 3.7.1.

We therefore compute the approximate gradient

$$[\boldsymbol{\nabla} S(\boldsymbol{\mu})]_i \approx \frac{1}{\sigma_{\boldsymbol{\mu};q}} \sum_{q=1}^{Q} \frac{1}{\operatorname{card}(\mathbb{D}_q)} \sum_{\mathbf{k} \in \mathbb{D}_q} \left(\frac{p_{\boldsymbol{\mu};q}(\mathbf{k}) - P_{\boldsymbol{\mu};q}}{\sigma_{\boldsymbol{\mu};q}} - \frac{p_q(\mathbf{k}) - P_q}{\sigma_q} \right) \frac{\partial p_{\boldsymbol{\mu};q}(\mathbf{k})}{\partial \mu_i}, \tag{4.23}$$

and the approximate Hessian

$$[\boldsymbol{\nabla}^2 S(\boldsymbol{\mu})]_{i,j} \approx \frac{1}{\sigma^2_{\boldsymbol{\mu};q}} \sum_{q=1}^{Q} \frac{1}{\operatorname{card}(\mathbb{D}_q)} \sum_{\mathbf{k} \in \mathbb{D}_q} \frac{\partial p_{\boldsymbol{\mu};q}(\mathbf{k})}{\partial \mu_i} \frac{\partial p_{\boldsymbol{\mu};q}(\mathbf{k})}{\partial \mu_j}, \tag{4.24}$$

where $i, j = 1, 2, \ldots, 6$.

4.3 CT/C-arm data

This section reflects the work of our collaborators from the M.E. Müller Institute for Biomechanics, Bern, Switzerland. The CT/C-arm data described in this section are used in Section 4.4.3 for validating the space-based algorithm.

4.3.1 Specimen preparation

A cadaver spine specimen was frozen so that it can be treated as a rigid body. Five fiducial markers (custom-made, titanium) were implanted on it. One was placed in the L5, two in the L4, and two in the L3 vertebra (Figure 4.10).

4.3.2 CT data acquisition

The specimen, placed on a plastic bag, was CT-scanned (using a GE LightSpeed Ultra CT scanner shown in Figure 4.1) with 72 slices of size 512×512 pixel each (Figure 4.11). The intra-slice pixel size was 0.36×0.36 mm, and the inter-slice thickness was 2.5 mm. The tilt angle was zero.

4.3.3 Tracking device

The position of all devices and instruments, equipped with infrared LED markers, was tracked using an optoelectronic position sensor (Northern Digital Optotrak 3020 shown in Figure 4.5).

4.3.4 Tracking of the fiducial markers

The coordinates of the fiducial markers were digitized in the R-COS using an optoelectronically tracked pointer.

4.3.5 C-arm geometric data acquisition

The images were acquired using a Siemens ISO-C C-arm (Figure 4.3) instrumented with LED markers. To get the geometric parameters of the C-arm x-ray projection, an approximation to the full perspective projection model was adopted by the authors of [43]. This approximation (an orthographic projection) was then used to determine an extrinsic, an intrinsic, and a mechanical calibration procedure that was subsequently applied to all C-arm images.

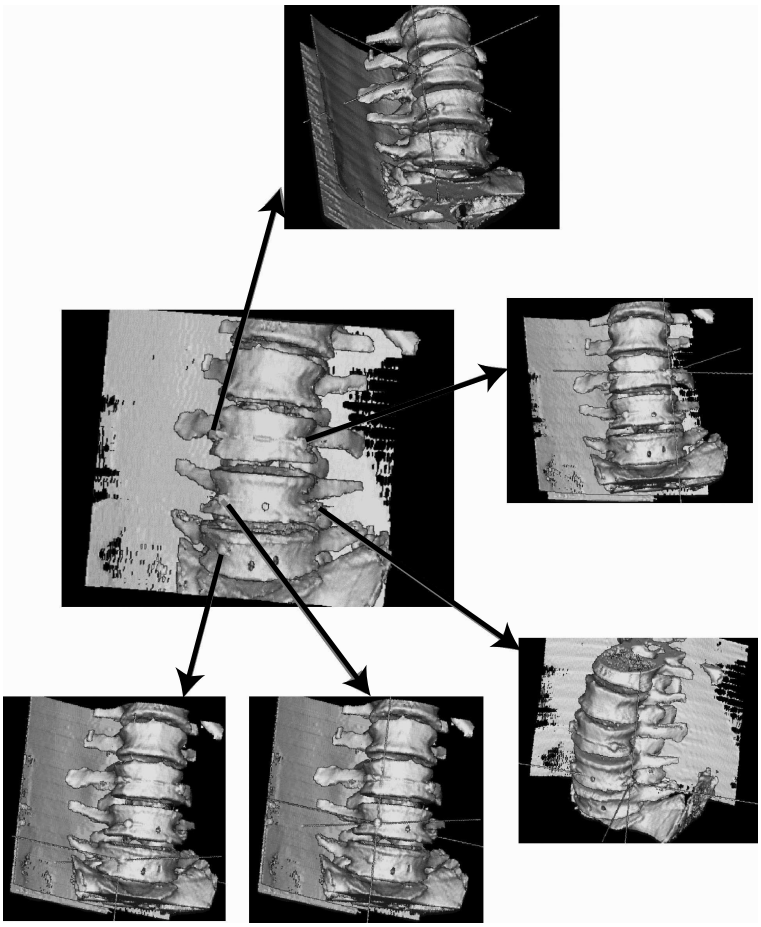


Figure 4.10: Position of the fiducial markers in CT data coordinate system. Courtesy of the M.E. Müller Institute for Biomechanics, Bern, Switzerland.

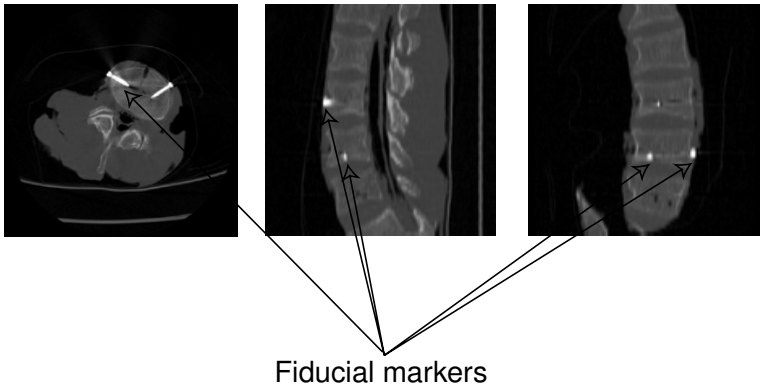


Figure 4.11: Transversal (left), sagittal (center), and frontal (right) CT slices with visible fiducial markers.

4.3.6 Navigation error

The optical tracking system provides data coordinates with a total RMS error of about 0.2 mm per LED marker. Imperfect calibration has a direct influence on the navigation error. The effect of all these errors on the overall accuracy with which a surgical tool is visualized in C-arm images was analyzed in [43].

Regarding the extrinsic calibration, a systematic error is introduced by the optoelectronic device. This produces a mean error of 0.4 mm for navigation within the volume between proximal and distal positions of the calibration plate. Residual errors of the mechanical calibration are estimated to have a maximum variation of 0.8 mm. Residual errors of the intrinsic calibration have a mean value of 1.2 pixel and a standard deviation of 1.0 pixel. The error associated to the use of an orthographic projection instead of a true perspective one is estimated to be at most 1.4 mm.

The overall accuracy study was performed using an optoelectronically tracked phantom. A navigation error with a mean of 0.55 mm, a standard deviation of 0.47 mm, and a maximum value of 2.34 mm was reported in [43].

4.3.7 Pixel size

The pixel size depends on the position of the image plane relative to the image intensifier. This position was chosen to be the proximal position so that it was possible to compute the A-COS coordinates of the image-plane corners based on

a known transformation from the I-COS indexes to the corresponding A-COS coordinates of the points belonging to the proximal plane. This transformation is known from the extrinsic calibration. Given the coordinates of the three image corners and the number of pixels along rows and columns of the image, the pixel size was computed.

4.3.8 C-arm image acquisition

A Dynamic Reference Base (DRB) was attached to two different positions on the specimen before shooting the C-arm images. The DRB defines the P-COS coordinate system. One set of images was shot with the DRB attached to one end of the spine specimen so that it is not visible in these C-arm images (for examples, see Figure 4.12). Another set was shot with the DRB attached to one of the vertebræ so that it is visible in these C-arm images (for examples, see Figure 4.13). The C-arm images were captured with seven different poses for each of the two sets. The image size was 768×576 pixel and the pixel size was 0.36×0.36 mm. The images were un-distorted before their further use.

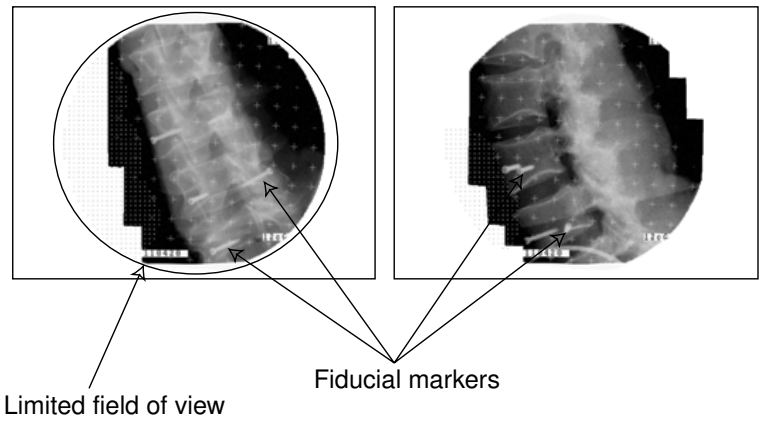


Figure 4.12: Two C-arm images where the DRB lies outside of the field of view.

The presence of the DRB in the C-arm images is a challenge to registration algorithms because it is not present in the CT data. This situation would be typical in a clinical setting.

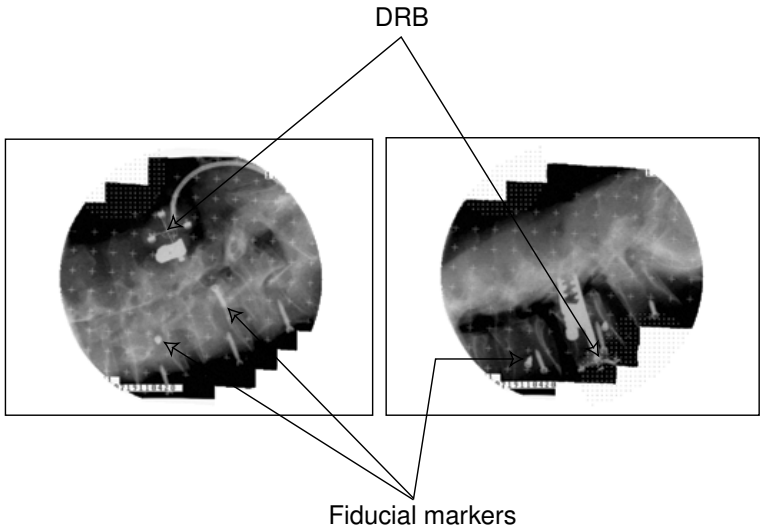


Figure 4.13: Two C-arm images with a visible DRB.

4.4 Validation

In this section, we validate our space-based algorithm using data with an *a priori* known ground-truth data alignment, which means that we can have an objective evaluation of the registration accuracy. For this purpose, we have designed a measure of misregistration that we introduce in Section 4.4.1.

In Section 4.4.2, we validate the algorithm in a fully controlled simulation environment. This section is based on our article [46]. In Section 4.4.3, we validate the algorithm using real CT/C-arm data based on fiducial markers. In this case, the true data alignment is known, albeit within some non-zero tolerance. This section is based on our article [47]. We discuss the validation results in Section 4.4.4.

4.4.1 Measure of misregistration

We transform every 3D CT index $\mathbf{n}$ into an R-COS coordinate by using two transformations: the ground-truth transformation B and the transformation $B_{\boldsymbol{\mu}}$ that we estimate by performing the volume-to-image registration. We define the misregistration as the average of the norm of the difference between the two

R-COS coordinates over all CT indexes, that is,

$$\varpi = \frac{1}{\mathrm{card}(f)} \sum_{\mathbf{n} \in f} \|(B - B_{\boldsymbol{\mu}})\, \mathbf{n}\|. \tag{4.25}$$

In case of perfect registration, which is achieved when $B_{\boldsymbol{\mu}} = B$, we have that $\varpi = 0$.

4.4.2 Case 1: Validation with synthetic data

Ground-truth registration

The most convenient way to have a complete and precise control over the geometric aspects of the experiment is that the C-arm images be obtained in the same fashion as we obtain the simulated ones. If we project the same volume data to compute the C-arm images and their simulations, we are potentially able to achieve perfect registration. To make the situation less ideal, we apply a known rigid-body geometric transformation $B_{\boldsymbol{\mu}_R}$ to the given volume f, which results in $f'(\mathbf{n}) = f(B_{\boldsymbol{\mu}_R}\mathbf{n})$. This operation introduces interpolation and clipping, which ensures that it is not possible anymore to find registration parameters $\boldsymbol{\mu}$ such that $S(\boldsymbol{\mu}) = 0$. Let us refer to the computed projections of the volume f' that has an arbitrary pose $\boldsymbol{\mu}'$ as the C-arm images, and to the computed projections of the volume f at pose $\boldsymbol{\mu}$ as to the simulated C-arm images. Figure 4.14 depicts the general test procedure. The goal of the registration is to find the parameters $\boldsymbol{\mu}$ for which the simulations are the most similar to the C-arm images, on average. The ground-truth alignment of the volume f to the C-arm images is therefore given by $B = B_{\boldsymbol{\mu}'} B_{\boldsymbol{\mu}_R}$.

Experiments

We performed experiments on a brain CT volume of size $32 \times 32 \times 32$ voxel (volume f). We synthesized two C-arm images of size 100×100 pixel whose projection planes were perpendicular to each other. We run the registration 100 times at full data resolution. For each registration, we generated a random initial value for the parameters $\boldsymbol{\mu}$. The misregistration was computed over 40 iterations in each case.

We have chosen $\boldsymbol{\mu}_R$ such that it translates the volume f by 1.2, 1.3, and 1.4 voxel along the x-, y-, and z-axes, respectively, and rotates it for about 6.6 degree. We have also selected $\boldsymbol{\mu}'$ such that it translates the volume f' by 10, 5,

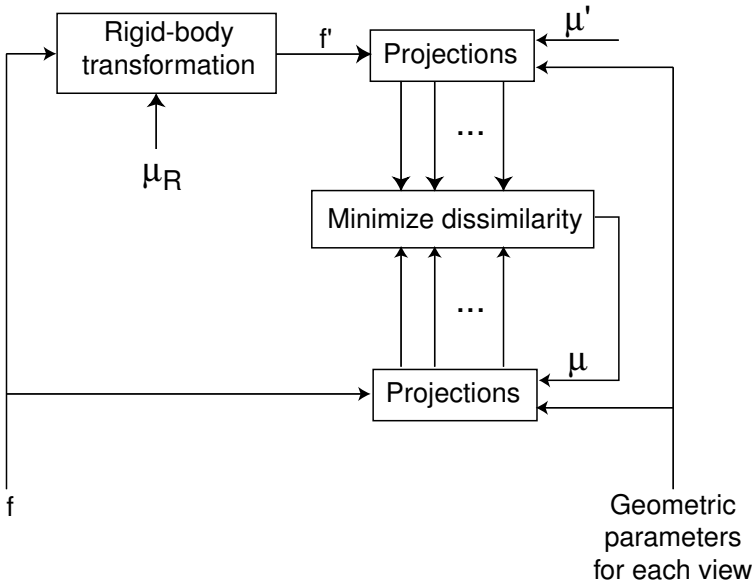


Figure 4.14: Diagram of the test procedure. Because of interpolation and clipping, there exists no pose $\boldsymbol{\mu}$ such that a projection of f is exactly identical to a projection of f'. This makes the registration task more challenging.

and 7 voxel along the x-, y-, and z-axes, respectively, and rotates it for about 93 degree.

The initial parameters $\boldsymbol{\mu}$ were chosen as consisting in part of a normally-distributed translation with mean 10, 5, and 7 voxel along the x-, y-, and z-axes, respectively, and standard deviation 1.8 voxel, and in part of a random rotation with mean 93 degree and standard deviation 2 degree.

The algorithm was successful and resulted in a subvoxel misregistration in each of the 100 cases. Starting from a large initial geometrical error ($\varpi = 4.79 \pm 0.71$ voxel),the algorithm succeeds to register the two images with $\varpi = 0.09 \pm 0.02$ voxel. Figure 4.15 shows ϖ in one of the 100 cases (solid line) on a linear-log scale. Between the first and the third iteration, the gradient algorithm dominates so that ϖ decreases slowly from iteration to iteration, by less than 25 percent. From the fourth iteration, the Newton algorithm dominates so that the amelioration of the fit accelerates, ϖ decreases by more than 50 percent

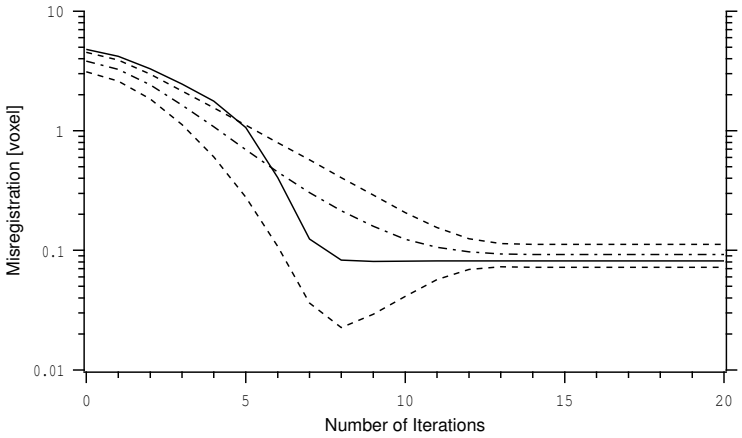


Figure 4.15: Performance of the registration algorithm for 100 randomly generated disturbances around the optimum solution. Solid line: Typical misregistration *vs.* number of iterations. Dash-dotted line: Average misregistration. Dashed lines: Average $\pm\frac{1}{2}$ standard deviation of the misregistration.

per iteration. After the ninth iteration, we observe that the algorithm has converged.

4.4.3 Case 2: Validation based on fiducials with real data

This validation is performed using the CT/C-arm data described in Section 4.3. To illustrate our method for simulating C-arm images (Section 4.2.4), we show two projections of the CT volume in Figure 4.16. These projections were computed for the volume pose at which the volume is in "almost" perfect registration with the two C-arm images shown in Figure 4.12.

Ground-truth registration

Having estimated the registration parameters $\boldsymbol{\mu}$, we evaluate (4.25) in three ways. In each case, we replace $B_{\boldsymbol{\mu}}$ by the transformation computed for the estimated parameters $\boldsymbol{\mu}$.

To measure the **registration accuracy**, we replace B in (4.25) by the transformation estimated using a given list of coordinates of the fiducial markers in the R-COS $(\mathbf{v}_i, i = 1, ..., 5)$ and in the CT volume $(\mathbf{n}_i, i = 1, ..., 5)$. The

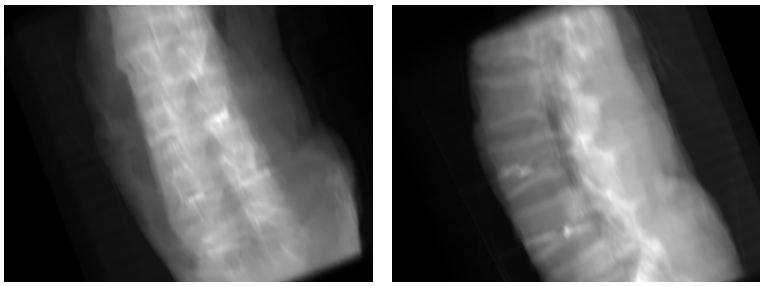


Figure 4.16: Simulated C-arm images that correspond to the C-arm images shown in Figure 4.12.

estimation is done by minimizing $\chi^2 = \frac{1}{5}\sum_{i=1}^{5} \|\mathbf{v}_i - B\,\mathbf{n}_i\|^2$ in terms of B. We denote by ϖ_a the corresponding value of ϖ.

To measure the **registration reproducibility**, we replace B in (4.25) by the transformation computed for the parameters $\boldsymbol{\mu}$ averaged over a set of experiments performed using the same data but different initial conditions. We denote by ϖ_r the corresponding value of ϖ.

To measure the **registration consistency**, we replace B in (4.25) by the transformation computed for the parameters averaged over the experiments on different data sets when using the same initial conditions. We denote by ϖ_c the corresponding value of ϖ.

Experiments

We performed experiments on the challenging data set, that is, using the C-arm images with a visible DRB. We use four-level CT-volume and C-arm-image pyramids (Table 4.1). These pyramids are dyadic in only two directions; we do not change the number of CT slices while performing the data reduction because the nominal inter-slice resolution is much coarser than the intra-slice one. We perform the registration at the two coarsest pyramid levels only, since this strategy gives a good tradeoff between accuracy and time. If we perform additional processing at the two finest data resolutions, we do not gain more than 0.2 mm accuracy. This does not justify the huge additional registration time. We could have coarsened the data further, for example by computing five-level pyramids, or by reducing the number of CT slices. However, such pyramids do not improve the performance of our registration algorithm since they result in either failures at the coarser data resolutions, or fail to perform

better at finer data resolutions.

Table 4.1: Four-level CT-volume and C-arm-image pyramids.

Level	CT Size [vxl]	C-Arm Image Size [pxl]
3	$64 \times 64 \times 72$	96×72
2	$128 \times 128 \times 72$	192×144
1	$256 \times 256 \times 72$	384×288
0	$512 \times 512 \times 72$	768×576

We run the algorithm with at least two C-arm images as input. One of the elements that determines the working range of our algorithm is given in terms of the angle between the two image planes. We have found that it operates safely if this angle is larger than 10°.

We performed three sets of experiments. The first set tests the performance of the algorithm when registering the CT to a single pair of C-arm images with the angle between their planes being about 90°, and when using different initial conditions. The second set registers the CT to several pairs of C-arm images using the same initial conditions for each image pair. These experiments aim at discovering how different values for the angle between two image planes influence the registration performance. The third set aims at finding out whether the algorithm performs better if it is provided with more than two input C-arm images; in each experiment of this type, the algorithm uses the same initial conditions.

Single image-pair experiments. We initialized the algorithm with 50 uniformly distributed random parameters $\boldsymbol{\mu}$ such that the initial values for ϖ_a and ϖ_r were $\varpi_a = 9.28 \pm 1.06$ mm and $\varpi_r = 9.48 \pm 1.28$ mm, respectively. The accuracy of the achieved registration was $\varpi_a = 1.47 \pm 0.0002$ mm, while its reproducibility was $\varpi_r = 0.003 \pm 0.001$ mm.

Several image-pairs experiments. Although we can make 21 image pairs out of the seven images that are at our disposal ($\frac{7!}{2!5!} = 21$), we performed experiments on only those pairs where the angle between the planes is larger than 10°. The number of such image pairs was 19.

After having imposed a deliberate initial misregistration such that the values for ϖ_a and ϖ_c were $\varpi_a = 9.02$ mm and $\varpi_c = 9.03$ mm, respectively, our

	Accuracy ϖ_a		Consistency ϖ_c	
	Mean [mm]	Deviation [mm]	Mean [mm]	Deviation [mm]
Initial	9.02	0.00	9.03	0.00
Level 3	2.10	0.69	1.83	0.90
Level 2	1.79	0.69	1.14	0.73
Level 1	unnecessary	unnecessary	unnecessary	unnecessary
Level 0	unnecessary	unnecessary	unnecessary	unnecessary

Table 4.2: Accuracy and consistency of the registration of the CT volume to 19 C-arm image pairs with the angle between any two image planes larger than 10°.

algorithm performed the registration with an accuracy $\varpi_a = 1.79 \pm 0.69$ mm (Table 4.2). The consistency of the registration was $\varpi_c = 1.14 \pm 0.73$ mm (Table 4.2).

Several-views experiments. We compared the quality of the registration based on two, three, four, and seven imaging views. We created three sets of C-arm images: image pairs, image triples, and image quadruples. The members of each set were chosen such that the angle between any two image planes in the set was larger than 15°. We thus performed the experiments on 18 image pairs out of 21, on 22 image triples out of 35 ($\frac{7!}{3!4!} = 35$), and on 13 image quadruples out of 35 ($\frac{7!}{4!3!} = 35$). To provide a fair comparison, we initialized the algorithm by the mean of the estimated parameters from the 19 image-pairs experiments. We were therefore initially close to the true solution. The initial data misregistration was $\varpi_a = 1.36$ mm.

The algorithm achieved the data registration with the accuracy and the consistency shown in Table 4.3 for the two-views experiments, in Table 4.4 for the three-views experiments, and in Table 4.5 for the four-views experiments, where it can be seen that an increase in the number of input images from two to three yields a much more accurate and consistent registration (with 95% confidence). On the contrary, the registration based on four input images is more consistent but not significantly more accurate (with 95% confidence) than the registration based on three images. The accuracy of the alignment by the parameters obtained in a seven-views experiment was $\varpi_a = 1.41$ mm, which belongs to the 95% confidence interval for the mean accuracy achieved in the

Table 4.3: Accuracy and consistency of the registration of the CT volume to 18 pairs of C-arm images with the angle between any two image planes larger than 15°.

	Accuracy ϖ_a		Consistency ϖ_c	
	Mean [mm]	Deviation [mm]	Mean [mm]	Deviation [mm]
Initial	1.36	0.00	0.12	0.00
Level 3	1.78	0.51	1.43	0.71
Level 2	1.66	0.49	0.99	0.61
Level 1	unnecessary	unnecessary	unnecessary	unnecessary
Level 0	unnecessary	unnecessary	unnecessary	unnecessary

Table 4.4: Accuracy and consistency of the registration of the CT volume to 22 triples of C-arm images with the angle between any two image planes larger than 15°.

	Accuracy ϖ_a		Consistency ϖ_c	
	Mean [mm]	Deviation [mm]	Mean [mm]	Deviation [mm]
Initial	1.36	0.00	0.32	0.00
Level 3	1.39	0.29	0.84	0.46
Level 2	1.39	0.24	0.48	0.22
Level 1	unnecessary	unnecessary	unnecessary	unnecessary
Level 0	unnecessary	unnecessary	unnecessary	unnecessary

three- and four-views experiments.

We performed a principal component analysis to illustrate the registration accuracy and consistency. We projected the estimated angles and translations onto their respective first principal components in Figure 4.17 for the two-views experiments, in Figure 4.18 for the three-views experiments, and in Figure 4.19 for the four-views experiments. We also give the projection of the solution determined by the fiducial markers, the mean of the parameters estimated in the corresponding experiment, and the parameters estimated in the seven-views experiment.

Table 4.5: Accuracy and consistency of the registration of the CT volume to 13 quadruples of C-arm images with the angle between any two image planes larger than 15°.

	Accuracy ϖ_a		Consistency ϖ_c	
	Mean [mm]	Deviation [mm]	Mean [mm]	Deviation [mm]
Initial	1.36	0.00	0.33	0.00
Level 3	1.28	0.23	0.51	0.29
Level 2	1.36	0.15	0.27	0.14
Level 1	unnecessary	unnecessary	unnecessary	unnecessary
Level 0	unnecessary	unnecessary	unnecessary	unnecessary

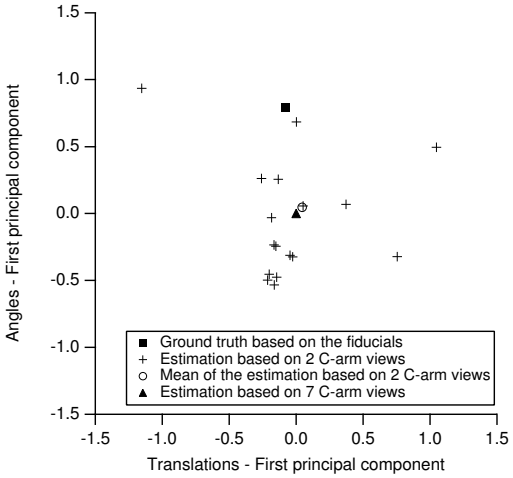


Figure 4.17: Projection onto the first principal component of the angles and of the translations estimated by registering the CT volume to 18 pairs of C-arm images with the angle between any two image planes larger than 15°.

Speed. As far as speed is concerned, the experiments were performed on a Power Mac G4, 733MHz, where the computation of volume and image pyramids took approximately 20 sec and 6 sec, respectively. We achieved the registration

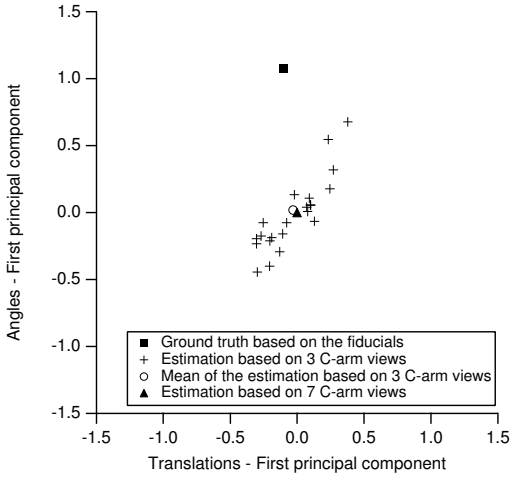


Figure 4.18: Projection onto the first principal component of the angles and of the translations estimated by registering the CT volume to 22 triples of C-arm images with the angle between any two image planes larger than 15°.

in 4 to 6 min (Table 4.6).

Table 4.6: Time required for the registration of the CT volume to pairs, triples, and quadruples of C-arm images.

Level	Two views		Three views		Four views	
	Mean [sec]	Deviation [sec]	Mean [sec]	Deviation [sec]	Mean [sec]	Deviation [sec]
3	50	11	70	12	82	21
2	128	66	179	86	173	71
1	ignored	ignored	ignored	ignored	ignored	ignored
0	ignored	ignored	ignored	ignored	ignored	ignored

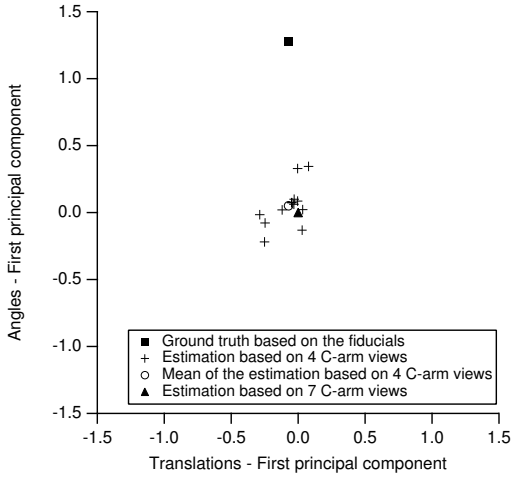


Figure 4.19: Projection onto the first principal component of the angles and of the translations estimated by registering the CT volume to 13 quadruples of C-arm images with the angle between any two image planes larger than 15°.

4.4.4 Discussion

In the experiments carried out with synthetic C-arm images, our algorithm registers pairs of images to the CT volume with a subvoxel accuracy of about 0.1 voxel. Part of this residual misalignment can be attributed to the transformation $B_{\boldsymbol{\mu}_R}$ which was introduced on purpose to make the registration task more challenging.

In the experiments performed with real C-arm images, our algorithm registers pairs of images to the CT volume with an accuracy of about 1.7 ± 0.5 mm and with a good reproducibility of less than 0.01 mm. The non-zero misregistration can be explained by the fact that some errors were made when digitizing the fiducial markers (the mean navigation error was 0.5 mm), and that some more errors were made when determining the CT indexes of the markers. An independent analysis has shown that the maximum error of determining the CT indexes of the markers was about 0.6 mm (0.25 slice off). This means that, in the experiments carried out, our algorithm is perhaps responsible for a lesser misregistration than that reported. Note that the mean misregistration of 1.7 mm is already subvoxel with respect to the inter-slice CT thickness of 2.5 mm.

We improve the registration accuracy and consistency when we augment the

number of input C-arm imaging views. These images should satisfy the working condition of the algorithm whereby the angle between at least two image planes must be larger than 10°. The mean registration quality significantly improves (with 95% confidence) if we perform the registration with three instead of two C-arm imaging views. In the experiments performed here, this improvement of the mean registration accuracy is about 0.3 mm (16%). The mean registration consistency improves by about 51%. The increase in the number of views from two to three slows down the algorithm for at most 1.5 min. Although the registration accuracy stays essentially the same, the registration consistency improves significantly (with 95% confidence) if we provide the algorithm with more than three C-arm imaging views. In the experiments performed, the mean registration consistency improves by about 44% when we switch from three to four input C-arm imaging views. The effect of the increase in the number of input C-arm images on the registration consistency can be seen as tightening the cloud of estimates around the seven-views estimate (Figures 4.17, 4.18, and 4.19).

The registration accuracy achieved using our method is consistent with the accuracy of similar approaches in the field. For instance, using a CT volume with the inter-slice thickness of 1.5 mm, Weese et al. [116] achieve an accuracy of about 0.5° for the rotations, 0.5 mm for the shifts parallel to the projection plane, and 5-6 mm for the height above the projection plane. Tomaževič et al. [100] report a RMS geometrical error of about 0.5 mm and a maximum geometrical error of about 1.2 mm when using a CT volume with the inter-slice thickness of 1 mm. We conclude that our algorithm could perhaps achieve even more accurate registration if our CT would have a smaller inter-slice thickness.

Chapter 5

REGISTRATION OF A 3D MODEL OF A PARTICLE TO ITS CRYO-EM IMAGES

In this chapter, we show the application of our frequency-based algorithm in 3D cryo-EM. First, a short introduction to cryo-EM is given in Section 5.1. In the same section, we review the methods for the assignment of the particle poses and give an overview of the contribution of our approach.

We use the the frequency-based algorithm from Chapter 3 since the parallel-beam projection geometry fits perfectly the electron-microscopy registration problem. Some algorithmic details that were left out in Chapter 3 are presented in Section 5.2.

In Section 5.3, we validate the algorithm using a 3D protein model from the Protein Data Bank (PDB) (http://www.rcsb.org/pdb/) and synthesized electron-microscopy images of this protein. These images have a known ground-truth alignment with respect to the 3D model of the protein.

We also have at our disposal a set of experimental electron-microscopy images and a 3D model of a protein for which we do not know the ground-truth alignment between the volume and the images. The performance of the algorithm using such data is investigated in Section 5.4.

To initialize our algorithm, we apply one of the standard methods for refinement of a particle pose assignment. For this purpose, we compare three methods commonly used in 3D EM.

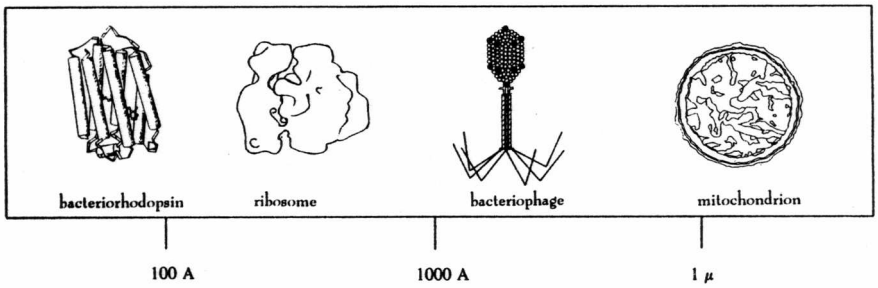


Figure 5.1: The range of biological structures covered by 3D electron microscopy. From Frank [26] (Fig. 1.1, p. 2).

5.1 3D electron cryo-microscopy

Cryo-EM [26, 50, 101] comprises methods that facilitate the visualization and the study of the structure of biological structures of size from 100 Å to 1 μ (Figure 5.1). These methods assume that an electron-microscopy image (electron micrograph) of a specimen is a 2D parallel-beam x-ray projection of the specimen. Its 3D structure (often called 3D map or 3D model) is computed from these projections at different viewing angles using image-reconstruction techniques suited to the nature of the specimen. This 3D model reveals not only the shape but also the interior density variations of the specimen at resolutions from molecular to near-atomic.

5.1.1 Nature of the sample

Different samples can be analyzed using cryo-EM. These are 2D crystals, helical-symmetry crystals, filamentous samples, single particles, macromolecular complexes, viruses, cell organelles, and bacterial cells.

A typical EM sample used for single-particle reconstruction contains many copies of the same object. These copies are assumed to have the same structure but completely unknown orientations.

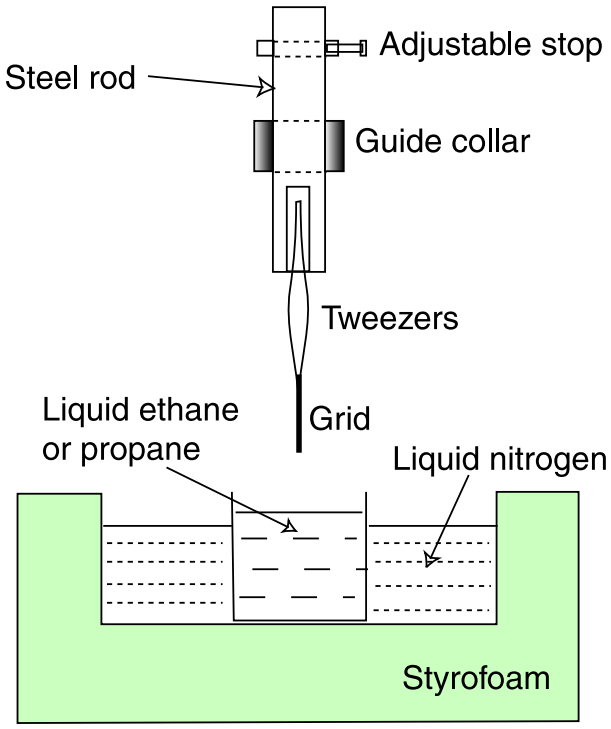


Figure 5.2: Vitrification set-up. From Frank [27].

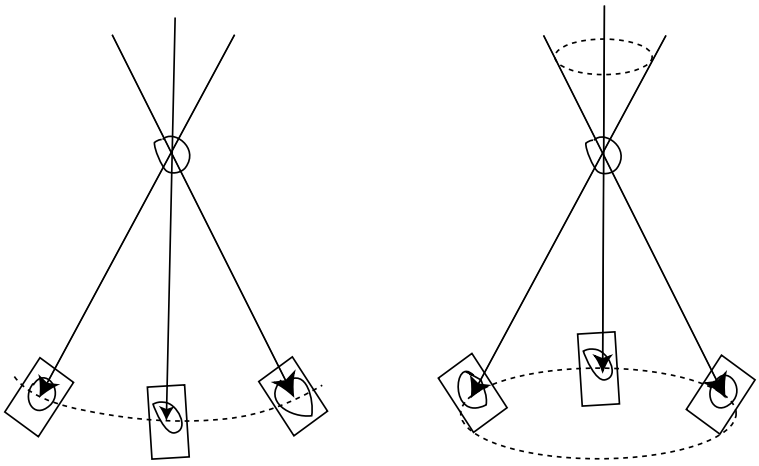


Figure 5.3: Single-axis and conical tilt collection geometry. From Frank [26].

5.1.2 Vitrification

The specimen is applied to an electron microscope grid within a thin water layer. The sample is rapidly frozen (vitrified) by plunging it into liquid ethane at liquid-nitogen temperature (Figure 5.2). Due to the rapid decrease in temperature, the water turns into vitreous ice that has properties akin to those of liquid water [27]. The most important property is that the water does not crystallize and that the molecules are not damaged.

5.1.3 Data collection

Electron diffraction data are recorded from large 2D crystals only. Images are recorded from non-tilted (all specimen types) and tilted samples (2D crystals, cell organelles, bacterial cells, certain single particles) [26, 101].

Four kinds of tilt-based data collection schemes are used: single-axis tilt, conical tilt, random-conical tilt, and general random tilt [26, 78]. Single-axis and conical tilt are shown in Figure 5.3. For a single-axis tilt series, the specimen is tilted in the microscope in a range of typically $-60°$ to $60°$ in small increments (e.g., $1°$ to $5°$), and an image is recorded for each position. For a conical tilt series, the specimen is tilted by one fixed angle in a range of $45°$ to $60°$. Then,

it is rotated within this plane by small angular increments (azimuthal angle), and an image is recorded for each position. These two schemes are mainly used for preparations that are radiation-resistant or that contain particles that have different shapes, that is, averaging over different particles is either not required or not possible. Random or random-conical tilt techniques intrinsically average over a large number of identical particles, and the electron exposure of each single particle can be kept very low. The random-conical scheme is based on the principle that many particles tend to exhibit a preferred orientation with respect to the specimen support plane, leaving only one rotational degree of freedom, which is a rotation around an axis perpendicular to the plane. The set of images from many such particles in a micrograph of a tilted (between 45° and 65°) specimen form a conical tilt series with random azimuthal angles. If a preferred orientation is not present, then any micrograph of the specimen showing particles in random orientation provides a random-tilt series.

There are two important properties of the acquired images. First, they are not equal to the theoretical projections of the specimen. This is mainly caused by the Contrast Transfer Function (CTF) of the electron microscope that we consider in Section 5.1.4. Second, they have a low signal-to-noise ratio (SNR) due to a limited radiation dose that may be delivered to the sample to avoid its destruction.

5.1.4 Contrast transfer function

In this section, we discuss the following: the origin of the CTF, the image formation model, the CTF model, the CTF estimation, and the CTF correction.

Image formation in the electron microscope is a complex process that relies on a multitude of interactions between the specimen and the electrons. It is therefore not surprising that the experimental projection is not equal to a theoretical, parallel-beam x-ray projection of the specimen. The differences between the two projections are caused by the electron source, magnetic lenses, defocus, and imaged matter. These differences are commonly modeled by the CTF which is a linear transfer function. The CTF theory is an approximation to a comprehensive theory of image formation that ignores many effects whose relative magnitudes vary from one specimen to the other [26].

Assuming that the projected 3D Coulomb potential distribution within the object ϕ is a real function, and that the noise is equal to zero, we write the FT of the image as

$$\hat{f}(\boldsymbol{\omega}) = \hat{h}(\boldsymbol{\omega})\hat{\phi}(\boldsymbol{\omega}), \tag{5.1}$$

where $\boldsymbol{\omega}$ is a spatial frequency, $\hat{\phi}$ is the FT of the projected electron density potential, and $\hat{h}$ is the CTF. In the space domain, the CTF is represented by the Point Spread Function (PSF) h that convolves ϕ to produce the image f, that is, $f(\mathbf{x}) = (h * \phi)(\mathbf{x})$, where $\mathbf{x}$ is the coordinate of a point in real space.

The PSF describes the way in which a single point of the object would be imaged by the electron microscope. The closer the PSF is to a delta function, the more faithful the image is to the theoretical projection of the object. However, the typical PSF is an oscillating function, with a central maximum that is often only slightly different from the surrounding maxima [26]. The CTF acts as a band-pass filter. Its typical consequence is that, in the case of a noisy micrograph, the particle is barely distinguishable from the background.

We use the model of image formation that has been proposed in [113]. It considers two types of noise: 1) noise before the image formation (n_b), that is, the background noise present in the image before it will be affected by the CTF, and 2) noise after the image formation (n_a), that is, the additive noise component due to the process of image recording and of image digitization. It is given by

$$\hat{f}(\boldsymbol{\omega}) = \hat{h}(\boldsymbol{\omega})(\hat{\phi}(\boldsymbol{\omega}) + \hat{n}_b(\boldsymbol{\omega})) + \hat{n}_a(\boldsymbol{\omega}). \tag{5.2}$$

Theoretical model of the CTF

A general model [26, 113] for $\hat{h}$ is given by

$$\hat{h}(\boldsymbol{\omega}) = K\,\hat{a}(\boldsymbol{\omega})(\sin(\hat{\chi}(\boldsymbol{\omega})) - \hat{q}(\boldsymbol{\omega})\cos(\hat{\chi}(\boldsymbol{\omega}))), \tag{5.3}$$

where $\hat{\chi}$ represents the shift of the phase of the scattered part of the electron wave behind the object, caused by the lens aberrations and the defocus. The term $\hat{q}$ is the ratio between the imaginary part and the real part of $\hat{\phi}$. As such, $\hat{q}$ is a function characteristic for each species, but within the frequency range of practical interest, it is safe to assume that it is constant, that is $\hat{q}(\boldsymbol{\omega}) = c_0$ [26]. The factor $\hat{a}$ is an attenuating envelope function that models the effects that result in a decrease of the CTF amplitude due to energy spread and partially coherent illumination, among others. The term K is a scaling factor.

The parameters of the CTF model proposed in [113] are: the voltage of the microscope, the chromatic aberration coefficient, spherical aberration coefficient, the maximum and the minimum defocus, the angle between the minimum defocus and the x-axis in the frequency domain, the energy spread of the electrons in the beam as a percentage of the voltage, the instability of the lens

Figure 5.4: Example of the estimated CTF. Courtesy of the Centro Nacional de Biotecnología, Madrid, Spain.

current as a percentage of the current, the semi-angle of the beam aperture, a drift of the sample that may produce mechanical displacements, a possible mechanical displacement, K, and c_0.

Estimation of the CTF

Many manual and automated techniques for determining the CTF have been proposed [26]. The automatic methods include techniques for fitting the theoretical parametric model of the CTF to the average power spectrum (periodogram) computed from the experimental images [26]. A more accurate CTF estimation [113] has been obtained by matching the theoretical model of the CTF to its Auto-Regressive Moving Average (ARMA) model computed from the experimental images.

Here, we often use the CTF estimated by the ARMA method [113]. This CTF is shown in Figure 5.4.

Correction of the CTF

Many attempts have been made to correct the CTF. They can be classified into three groups.

The first group consists of methods for instrumental correction of the CTF. Their goal is to improve the electron microscope so that the PSF behaves more like a delta function [26].

The second group contains the techniques for computational correction of the CTF that require an estimate of the CTF. The simplest correction of the CTF is the phase correction of $\hat{f}$. This is achieved by reversing the sign of the phase of the image at the frequencies where the CTF is negative [26]. A method for the CTF correction that corrects the amplitude and the phase of the image at the same time is Wiener filtering [26].

In the third group, we can put the techniques that do not correct the CTF explicitly but rather that tend to improve the techniques of specimen preparation and the imaging conditions. For instance, S. De Carlo et al. [22], showed that the technique of cryo-negative staining increases the image contrast when compared to the unstained technique. Also, one can use a much higher defocus to make the particle more visible in the image [26].

We use the technique of phase correction for the experiments with synthetic data. In contrast, we perform the real-data experiments with the CTF-uncorrected data obtained using the cryo-negative staining technique of S. De Carlo et al. [22].

5.1.5 Particle reconstruction

Every technique for particle reconstruction consists of two steps. First, the poses of the particle images are determined [21, 30, 32, 37, 52, 73, 79, 93, 111]. A review of the literature on the particle pose determination can be found in Section 5.1.6. Second, a reconstruction algorithm is applied to compute a 3D model of the particle using the particle images and their poses. A review of the 3D reconstruction methods can be found in Section 5.1.7. Sometimes, the two steps are performed simultaneously [77, 114].

Electron diffraction data are processed using a technique called electron crystallography. They provide the highest resolution 3D models (3-10 Å).

Electron tomography is a technique used for particle reconstruction from images of tilted samples. It yields 3D models of a lower, usually nanometer resolution. There are two reasons for that. One is a "missing cone" problem due to a limited maximum tilt angle in the electron microscope. The other is related to differences in the defocus within the tilted image.

Single-particle techniques are suitable in case of large macromolecular complexes that often resist crystallization or allow to be only partially crystallized, after removing their flexible parts [27]. Using zero-tilt single-particle techniques, 3D reconstruction at resolutions of 5-15 Å has been achieved for a number of different specimens (icosahedral-symmetry viruses, D6 pointgroup-symmetry hemoglobin, asymmetric E. coli ribosome) [111].

Here, we restrict ourself to single-particle reconstruction.

5.1.6 Step 1: Determination of the particle poses

There are two groups of techniques for determining the pose of the single particles. The first family does not require a reference 3D model at the input. The

method of moments and the common-line technique (known also as angular reconstitution, or geometrical method) belong to this group. The other group relies on the reference model and is supposed to refine the angular assignment in order to get a finer model. There are two types of reference-model-based techniques. One performs the assignment in a continuous-parameter space while the other does it in a discrete-parameter space. The algorithms developed in this thesis assign the pose in a continuous-parameter space.

In this section, we review the methods based on moments, common lines, and a reference model (with discrete and continuous parameters). We also point to the novelties of our approach.

Method of moments

This method uses a known relationship between the area moments of the 3D object and the moments of its 2D projections [5, 6, 30, 32, 85, 115]. The method based on low-order moments is fast but not precise. The use of high-order moments is not recommended since they are sensitive to noise and may therefore yield unstable results.

Common-line method

This method involves a classification of the single-particle images and a class averaging to reduce noise [111]. The geometrical relationships between the class averages, in terms of the two out-of-plane rotations, are computed from the angles between 1D line projections that 2D projections of a 3D object have in common [32, 73, 111]. The common-line projection for the two 2D projections is found by comparing their sinograms line-by-line, and by identifying the maximum of the sinogram correlation. At least three different projections are required to orient the images of an entirely asymmetric particle.

The 1D FT of the line projection that two different 2D projections of a 3D object have in common corresponds to a central line that the 2D FTs of the 2D projections have in common. A FT-domain equivalent of the real-space common-line projection method can be found in [32]. A technique that extends upon the FT-domain common lines and focuses on dealing with translational displacements and with noise has been proposed in [52].

Discrete-parameter method

There are two types of methods that determine the particle pose by a set of discrete parameters. One is based on a library of projections of the reference

volume covering the space evenly [73, 93]. The other relies on the 3D Radon Transform (RT) of the volume (sinogram) [79].

In [73, 93], volume projections are computed by covering the entire range of the two out-of-plane rotation angles. Penczek et al. [73] sample the volume projections in polar coordinates and compute the FTs of all the rings. They sample the centered particle images in polar coordinates and compute the FTs of all the rings. They compute 1D cross-correlation functions in polar coordinates between the particle images and all the projections in the library. The two out-of-plane rotation angles are determined by the largest cross-correlation coefficient. The in-plane rotation angle is determined by the position of the maximum in the corresponding cross-correlation function. This method is available in the SPIDER package [28]. We will refer to it as space-projection matching.

For each reasonable translation and rotation of the particle image, Sorzano et al. [93] search for the projection in the library that best matches it. They select the best-matching projection with the highest correlation coefficient. The pose of the particle image is determined by the out-of-plane rotation parameters of this projection and the best-matching in-plane rotation and translation parameters of the particle image. The correlation is computed in coarse-to-fine fashion using a discrete wavelet transform. This reduces the computation complexity and increases robustness with respect to noise. This method is available in the Xmipp package [60]. We will refer to it as wavelet-projection matching.

Radermacher [79] computes the 2D RT of each image and the 3D RT of the volume. He then computes a cross-correlation that depends on five parameters: three angles and two translations. The parameters for which this function assumes its maximum define the pose of the particle image. The advantage of that method is that the 3D RT contains the same information as a discrete set of all possible volume projections but it is represented using a smaller data set. This method is available in the SPIDER package. We will refer to it as Radon angular assignment.

The discrete-parameter methods use exhaustive-search techniques to determine the optimal parameters. A general disadvantage of these methods is that their accuracy is affected by the step with which the parameters are discretized. The smaller is the step, the more accurate, but slower is the parameter assignment.

Continuous-parameter method

A method that relies on the CST-based least-squares fit of the 2D FT of each particle image into the 3D FT of the particle model has been proposed in

[77, 114]. It combines the information from a tilt series of a disordered collection of many identical particles to obtain a 3D reconstruction. It models the 3D FT of the object's electron density by a truncated expansion in a complete orthonormal set of basis functions in spherical polar coordinates. The data are compressed, and the expansion coefficients together with the parameters of particle orientations are estimated for all particles simultaneously. The estimation is performed by using a modified Gauss-Newton minimization of a weighted variance of the fit of the model to the compressed data.

Novelties that our algorithm brings. Our algorithm belongs to the continuous-parameter methods. It relies on an iterative minimization of a least-squares measure of dissimilarity between the 2D FT of the image and extracted central slices of the 3D FT of the volume. To extract the slices, we interpolate the 3D DFT of the volume using a cubic B-spline data model. To make the algorithm robust to noise and less sensitive to the quality of the initial estimate for the unknown parameters, we apply our multiresolution strategy based on the frequency-domain weighting of the cost function. We propose a fast initialization of our algorithm by a method that relies on a coarse library of volume projections (Figure 5.5). Our algorithm can be used in alternation with 3D reconstruction for further iterative refinements (Figure 5.6).

Our algorithm differs from the one described in [77, 114] in several ways. This previously published method optimizes a much larger number of parameters than ours, which means a higher risk for the algorithm to get trapped into local optima of the cost function, and an augmented dependence on the quality of the starting estimates for the parameters. In contrast, our algorithm uses a model that is created from electron micrographs of the object. In return, it requires interpolation of the 3D FT of the volume. Since it estimates 5 parameters per particle only, for each particle independently, it allows an easier parameter control. Our algorithm is versatile since it can be applied in combination with any available reconstruction algorithms within a procedure for iterative refinement of the estimated particle poses.

5.1.7 Step 2: 3D reconstruction

The reconstruction can be accomplished in real space or in the frequency domain. The methods that are based on the FT-CST and that perform the interpolation in the frequency domain are called Fourier reconstruction algorithms. The methods that are rely on the FT-CST and that use the interpolation in real space are called back-projection methods. The methods that refine the solution

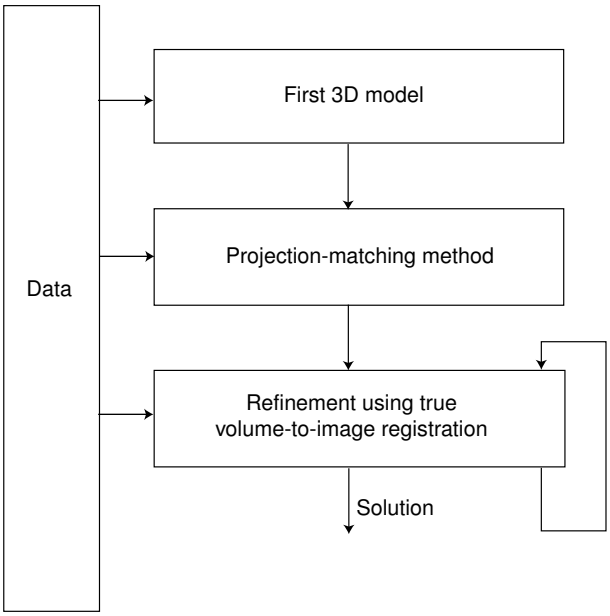


Figure 5.5: Our strategy for refining a 3D model of a particle.

in real space iteratively but that are not based on the CST are called Algebraic Reconstruction Techniques (ART).

Fourier reconstruction algorithms [26, 67, 68] assemble the 2D FTs of the object's projections at different angles into an estimate of the 3D FT of the object and then invert the 3D FT to obtain the estimate of the object. The position of the samples of the 3D FT of the object generally do not coincide with the regular 3D Fourier grid. The computation of the value of the 3D FT on grid points, given these samples at arbitrary positions, is a difficult interpolation problem.

Back-projection methods [48, 67, 68, 78] bring into play the FT-CST by rewriting the inverse FT in polar coordinates and rearranging the limits of the integration. The Weighted Back-Projection (WBP) method [78] consists of two steps: weighting of each image in the frequency domain by a weighting function (the filtering part) and summing the inverse FT of the weighted images (the backprojection part). These methods are usually more accurate than Fourier reconstruction methods since they perform interpolation in real space. Simple linear interpolation is usually satisfying in the case of a backprojection algo-

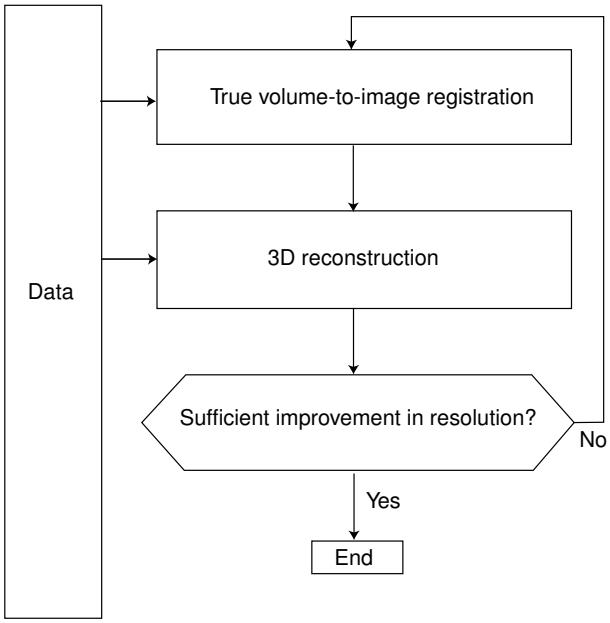


Figure 5.6: Further refinement of a 3D model of a particle using true volume-to-image registration.

rithm while more complicated interpolation schemes are required for the Fourier methods.

Since WBP methods can deal with arbitrary tilt geometry, they are currently the most widely used methods for reconstructing single particles from their transmission electron-microscopy images [26, 78]. They have been specifically developed for conical-tilt geometry. For single-axis tilt geometry, the WBP algorithm is equivalent to the standard Filtered Back-Projection (FBP) method. The FBP method uses a weighting function that is proportional to the frequency magnitude [48]. The WBP methods assume more complicated weighting functions for arbitrary tilt geometry [78].

To produce a reconstruction that is accurate and without artifacts, back-projection methods require a large number of images that are uniformly distributed in the 3D space. The requirement for providing an entire angular coverage becomes clear from the formulation of the reconstruction problem as data collection in the Fourier space. For example, the typical maximum tilt

angle of 60° for single-axis tilt geometry means that the data set will miss an entire range of central-plane orientations. This will produce a missing cone in the 3D Fourier space. The reconstruction coming from these data will show artifacts such as elongation of features in the direction of the missing angular range.

ART [48, 59, 67, 68] have been shown to be useful in overcoming the problem of missing angular range. ART is an iterative space-domain method that produces better-quality reconstructions at the expense of convergence.

In this work, we use zero-tilt single-particle images that are known to suffer from non-uniform angular sampling. To obtain good-quality reconstructions with such data, we take advantage of ART with smooth spherically symmetric volume elements [59]. This method is described next.

Algebraic Reconstruction Techniques

This is a series-expansion method. It assumes that the volume f to be reconstructed can be approximated by a linear combination of J basis functions $b_j, j = 1, \ldots, J$. Each of these functions is a known function b shifted to one of J grid points $\mathbf{g}_j$. That is,

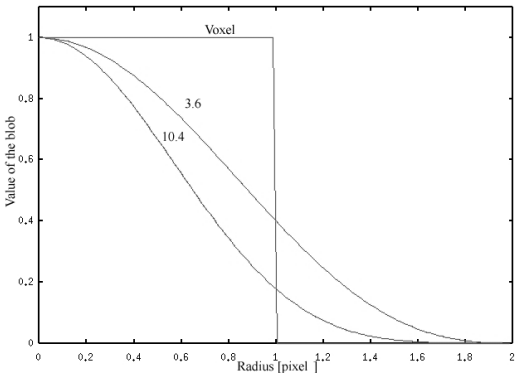
$$f(\mathbf{r}) \approx \sum_{j=1}^{J} c_j \, b_j(\mathbf{r}) = \sum_{j=1}^{J} c_j \, b(\mathbf{r} - \mathbf{g}_j). \tag{5.4}$$

Let us refer to a pixel value in the set of N experimental projections (each containing M pixels) as a measurement of the volume to be reconstructed. Let y_i be the i-th measurement, $i = 1, \ldots, M \times N$. Let $s_{i,j}$ be the corresponding line integral of b_j. Using the model (5.4), we have the following approximation

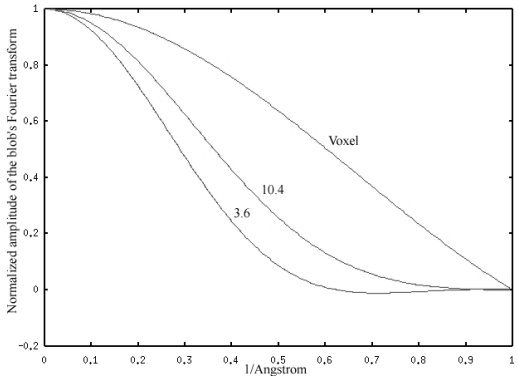
$$y_i \approx \sum_{j=1}^{J} c_j \, s_{i,j}. \tag{5.5}$$

The goal of the algorithm is to estimate the unknown reconstruction coefficients c_j from this system of approximate equations.

The simplest basis functions are cube-shaped elements that have the value 1 inside a single voxel and the value 0 everywhere else. In that case, the approximation (5.4) will have the value c_j inside the j-th voxel. More accurate models can be obtained using trilinear elements with a support extending over a cube of eight voxels. We use an ART algorithm with generalized Kaiser-Bessel window functions (blobs) [56, 57, 59]. These functions are spherically symmetric, spatially limited, and have a tunable smoothness.



(a) 1D blob profile.



(b) Fourier transform of the profile from Figure 5.7(a).

Figure 5.7: 1D blob profile and its Fourier transform for $a = 2, m = 2$, and two values of α together with a cube-shaped blob. From Sorzano [92].

A blob is determined by three parameters: the radius of the sphere within which the blob has a non-zero value (a), the number of times the blob is con-

tinuously differentiable (m), and the control of the density drop-off of the blob that can be used to insure that the blob's FT becomes approximately zero at a user-specified frequency (α). The blob's smoothness controls the smoothness of the reconstruction. Its spherical symmetry allows easy calculation of $s_{i,j}$. Blobs have been shown to perform better than voxels in reducing incorrect features and noise [56]. This is explained by their "almost" band-limitedness and by a smooth transition to zero at their spatial limit. For illustration purposes, we show in Figure 5.7 1D blob profiles and theirs FTs for $a = 2, m = 2$, and two values of α. A cube-shaped blob is also shown in these figures.

The arrangement of the centers of shifted versions of the basis functions is referred to as a grid. The quality of the reconstruction is affected by the choice of the grid as much as it is affected by the choice of the three blob parameters.

For the particle reconstruction, we use the ART algorithm based on the body-centered cubic grid (Figure 5.8) that has been found useful in electron-microscopy applications [59, 61, 62]. More precisely, we use the standard blobs ($a = 2, m = 2, \alpha = 10.4$) with the grid-point distance $g = \sqrt{2}$ [62].

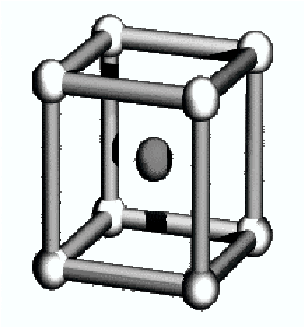


Figure 5.8: Body-centered cubic grid.

The particular variant of ART that we apply estimates the vector $\mathbf{c}$ (the J-dimensional vector of coefficients c_j) by refining an initial estimate [59]. We choose the initial estimate to be a zero vector. We use the data from only one projection per iteration. We cycle through all the projections, that is, the update of the estimate for the iteration $k + 1$, $(k = 0, \ldots, N - 1)$ is

$$\mathbf{c}^{(k+1)} = \mathbf{c}^{(k)} + \lambda \sum_{i=k\,M+1}^{(k+1)\,M} \frac{y_i - \langle \mathbf{s}_i, \mathbf{c}^{(k)} \rangle}{\|\mathbf{s}_i\|^2} \mathbf{s}_i, \tag{5.6}$$

where λ is a real number called the relaxation parameter that controls the magnitude of the update, $\mathbf{s}_i$ is the J-dimensional vector whose j-th component is $s_{i,j}$. The relaxation parameter controls the speed of the convergence and reduces the influence of the noise.

5.2 Registration algorithm

In this chapter, we assign the pose of a cryo-EM image with respect to a 3D particle model following the algorithm shown in Figure 5.9. The input data for the algorithm are: 1) the 3D model of the particle, 2) the cryo-EM images, 3) the pixel and voxel sizes, and 4) the initial volume pose.

The apodization of the input volume and images by Gaussian windows is represented by the "Windowing" blocks in Figure 5.9. The purpose of apodization is to enforce nearly space-limited data, which allows us to interpolate them in the frequency domain (Section 2.1.1).

In Chapter 3, we already defined the volume pose for the electron-microscopy application of our algorithm. Here, we apply the CST approach to the computation of the FTs of the volume projections as well as the multiresolution weighting of the cost function. In this section, we give the expressions for the multiresolution cost function, the gradient, and the Hessian.

Note that, in this application, we do not optimize the cost function in all six parameters of the volume pose but in only five of them. The unknown parameters are $\boldsymbol{\mu} = (\varphi, \theta, \psi, \Delta x, \Delta y)$. We cannot recover Δz because the parallel-beam projection geometry is invariant with respect to the shift along the projection direction (the z-axis here).

5.2.1 Multiresolution cost function

We minimize the following measure of dissimilarity between the FT of a cryo-EM image and the FTs of the volume projections:

$$S(\boldsymbol{\mu}) = \frac{1}{2} \sum_{\mathbf{n} \in \mathbb{Z}^2_*} w(\mathbf{n}) \left[\left(\frac{R_{\boldsymbol{\mu}}(\mathbf{n})}{\sigma_{\boldsymbol{\mu}}} - \frac{R(\mathbf{n})}{\sigma} \right)^2 + \left(\frac{I_{\boldsymbol{\mu}}(\mathbf{n})}{\sigma_{\boldsymbol{\mu}}} - \frac{I(\mathbf{n})}{\sigma} \right)^2 \right], \qquad (5.7)$$

where $\mathbb{Z}^2_* = \mathbb{Z}^2 \backslash \{\mathbf{0}\}$, $R_{\boldsymbol{\mu}}$ and $I_{\boldsymbol{\mu}}$ are the real part and the imaginary part of the 2D DFT of the projection of the volume at pose $\boldsymbol{\mu}$, respectively, and R and I are the real part and the imaginary part of the 2D DFT of the cryo-EM image, respectively. We do not compute the 2D DFTs of the projections explicitely.

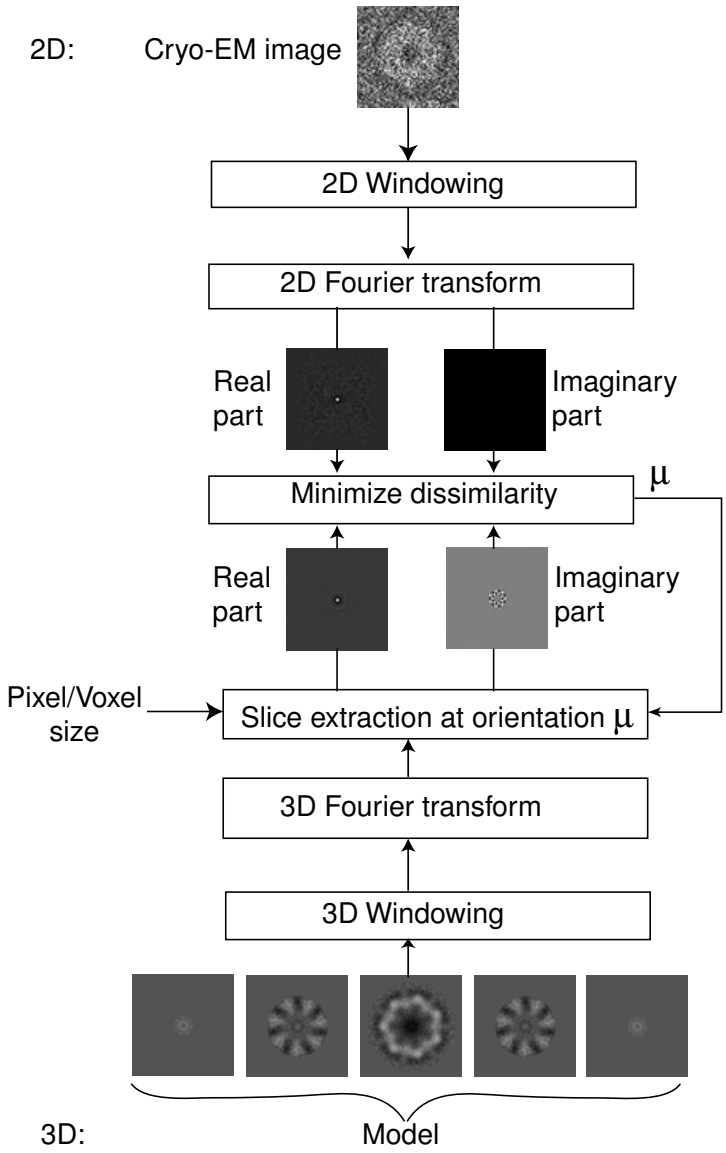


Figure 5.9: Frequency-based algorithm for assigning the pose of a cryo-EM image with respect to a 3D particle model.

We rather extract the corresponding central slice of the 3D DFT of the volume according to the CST approach described in Section 3.5.

Cryo-EM images and projections have different ranges of intensities. To take this into account, we normalize them by their standard deviations, σ and $\sigma_{\boldsymbol{\mu}}$, respectively, where $\sigma^2 = \sum_{\mathbf{n}\in\mathbb{Z}^2_*} (R^2(\mathbf{n}) + I^2(\mathbf{n}))$, and $\sigma^2_{\boldsymbol{\mu}} = \sum_{\mathbf{n}\in\mathbb{Z}^2_*} (R^2_{\boldsymbol{\mu}}(\mathbf{n}) + I^2_{\boldsymbol{\mu}}(\mathbf{n}))$. We implement the coarse-to-fine processing in the frequency domain by weighting the cost function by a 2D Gaussian-shape function w, as explained in Section 3.6.2.

5.2.2 Approximated gradient and Hessian

For simplicity (as in Chapter 4), we assume that the first derivative of $\sigma_{\boldsymbol{\mu}}$ with respect to $\boldsymbol{\mu}$ is equal to zero. The algorithm thus computes the approximate gradient

$$\begin{aligned}[\boldsymbol{\nabla} S(\boldsymbol{\mu})]_i \approx \tfrac{1}{\sigma_{\boldsymbol{\mu}}} \sum_{\mathbf{n}\in\mathbb{Z}^2_*} \quad & w(\mathbf{n}) \left(\tfrac{R_{\boldsymbol{\mu}}(\mathbf{n})}{\sigma_{\boldsymbol{\mu}}} - \tfrac{R(\mathbf{n})}{\sigma} \right) \tfrac{\partial R_{\boldsymbol{\mu}}(\mathbf{n})}{\partial \mu_i} + \\ & w(\mathbf{n}) \left(\tfrac{I_{\boldsymbol{\mu}}(\mathbf{n})}{\sigma_{\boldsymbol{\mu}}} - \tfrac{I(\mathbf{n})}{\sigma} \right) \tfrac{\partial I_{\boldsymbol{\mu}}(\mathbf{n})}{\partial \mu_i},\end{aligned} \tag{5.8}$$

where $i = 1, 2, \ldots, 5$.

For reasons given in Section 3.7.1, we ignore the second-order term in the true Hessian (3.31). The algorithm thus uses the approximate Hessian

$$\begin{aligned}[\boldsymbol{\nabla}^2 S(\boldsymbol{\mu})]_{i,j} \approx \tfrac{1}{\sigma^2_{\boldsymbol{\mu}}} \sum_{\mathbf{n}\in\mathbb{Z}^2_*} \quad & w(\mathbf{n}) \tfrac{\partial R_{\boldsymbol{\mu}}(\mathbf{n})}{\partial \mu_i} \tfrac{\partial R_{\boldsymbol{\mu}}(\mathbf{n})}{\partial \mu_j} + \\ & w(\mathbf{n}) \tfrac{\partial I_{\boldsymbol{\mu}}(\mathbf{n})}{\partial \mu_i} \tfrac{\partial I_{\boldsymbol{\mu}}(\mathbf{n})}{\partial \mu_j},\end{aligned} \tag{5.9}$$

where $i, j = 1, 2, \ldots, 5$.

5.3 Validation using data with a known ground-truth alignment

In this section, we validate our frequency-based algorithm in a fully controlled simulation environment using data with an *a priori* known ground-truth alignment. A known ground-truth alignment means that we can evaluate the registration accuracy objectively. For this purpose, we have designed a measure of misregistration that is presented in Section 5.3.1. As volumetric data, we used a 3D model of a protein from the PDB. As reference images, we used a set of simulated cryo-EM images obtained by projecting the 3D protein model

at known poses and applying noise and the CTF, subsequently. The volume data, the synthesized cryo-EM images, and the synthesizing procedure are described in Section 5.3.2. In Section 5.3.3, we show the ability of the frequency-based method to refine an initial pose assignment obtained applying one of the standard algorithms for the particle pose assignment. We also compare three standard algorithms in order to decide which one to use for the initialization.

The results are given in terms of misregistration. The histograms of the misregistration were computed with a bin width of 0.35 voxel.

5.3.1 Measure of misregistration

We transform the 3D index of each voxel of a virtual volume, given in homogeneous coordinates $\mathbf{n} = (n_1, n_2, n_3, 1)$ by using two transformations: the ground-truth transformation B and the estimated transformation $B^{\alpha}_{\boldsymbol{\mu}}$. Both transformation matrices, B and $B^{\alpha}_{\boldsymbol{\mu}}$, are of size 4×4 (homogeneous coordinates) and comprise the rotation, translation, and scale parameters. The transformation $B^{\alpha}_{\boldsymbol{\mu}}$ takes into account the symmetries of the particular protein. It involves a rotation by an angle α around the symmetry axis before applying the transformation parameters $\boldsymbol{\mu}$ estimated by performing the volume-to-image registration. In our work, we have mostly used proteins with a threefold rotational symmetry. In this case, α takes the following values: 60°, 120°, and 240°. We compute the average of the norm of the difference between the two transformed coordinates over all $\mathbf{n}$ for each α. We define the misregistration as the minimum averaged norm, that is,

$$\varpi = \min_{\alpha} \frac{1}{\mathrm{card}(f)} \sum_{\mathbf{n} \in f} \|(B - B^{\alpha}_{\boldsymbol{\mu}})\,\mathbf{n}\|. \tag{5.10}$$

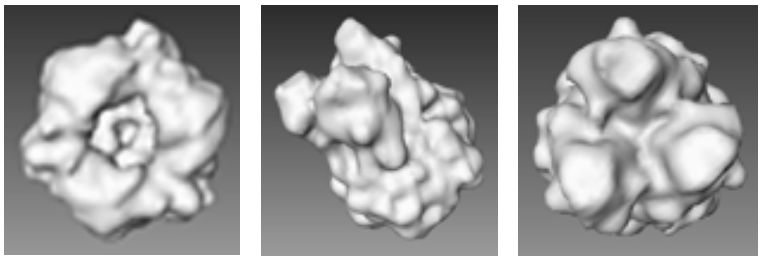
A smaller ϖ indicates a more accurate pose assignment. The perfect assignment should result in $\varpi = 0$.

5.3.2 Data

Volume

We used the Adenovirus Type 5 Hexon volume from the PDB (accession code: 1RUX [84]). The size of the volume was $64 \times 64 \times 64$ voxel and its sampling rate 3.27 Å.

The human adenoviruses are known to cause respiratory illness, conjunctivitis, and gastroenteritis. They are a potential cause of sexually transmitted



Three views of an isosurface.

A few slices.

Figure 5.10: 3D model of the Hexon from the PDB.

diseases. Adenovirus Type 5 causes mild respiratory infections. This type is mainly used for gene therapy. It consists of a shell that contains three proteins: hexon, penton base, and fiber. The hexon is the largest among the three proteins. The 3D model of the Hexon protein, used here (Figure 5.10), has been obtained by x-ray crystallography. This protein has a threefold rotational symmetry.

Synthetic cryo-EM images

These images were synthesized by C. Ó. S. Sorzano (Centro Nacional de Biotecnología, Madrid, Spain) in a two-step procedure as in [93].

In a first step, 1,000 projections of the reference model were computed with random poses determined by a set of uniformly distributed angles $\varphi \in [0°, 120°]$ and $\theta \in [0°, 175°]$. The three remaining parameters ψ, Δx and Δy were equal to zero for each projection.

In a second step, noise was added and the CTF was applied according to (5.2). Noise and the CTF were simulated using a model proposed by Velázquez-Muriel et al. [113]. The noise before the CTF (n_b) was modeled by a Gaussian white noise with the standard deviation $\sigma_b = 0.54$. The noise after the CTF

(n_a) was modeled by a low-pass filtered Gaussian noise with the standard deviation $\sigma_a = 2.20$ generated at the input $n_a^{'}$. The standard deviations were adjusted so that the ratio between the power of the signal and the power of the total noise was approximately equal to 1/3. The filter (Figure 5.11) and the CTF (Figure 5.4) have been estimated simultaneously [113] from an electron micrograph for the following parameters of the electron microscope: accelerating voltage=200 kV, defocus=-27700 Å , spherical aberration=2 mm, and convergence cone=0.21 mrad. The CTF, the background spectrum filter, and the synthesized cryo-EM images were of size 64×64 pixel. The synthesizing procedure is schematically shown in Figure 5.12. A few examples of these random projections and the corresponding synthesized cryo-EM images can be seen in Figure 5.13.

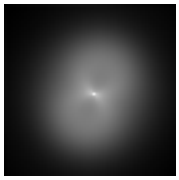


Figure 5.11: Low-pass filter used to simulate noise for the synthetic cryo-EM images. Courtesy of the Centro Nacional de Biotecnología, Madrid, Spain.

5.3.3 Experiments

Initialization

Three standard algorithms for the particle pose refinement were compared by C. Ó. S. Sorzano [93] using different proteins from the PDB. The algorithms under investigation were wavelet-projection matching, space-projection matching, and Radon angular assignment. The wavelet-projection matching and the space-projection matching are based on a library of projections of the reference volume. Each of the two libraries contained 500 images uniformly distributed in the 3D space with an angular step of 5°. The Radon transform was also sampled every 5°. We show here the results of this comparison for the Hexon data described in Section 5.3.2.

According to the mean value and standard deviation of the misregistration (Table 5.1), the wavelet-projection matching was the most accurate (with 95% confidence) among the three methods. We have therefore decided to initialize

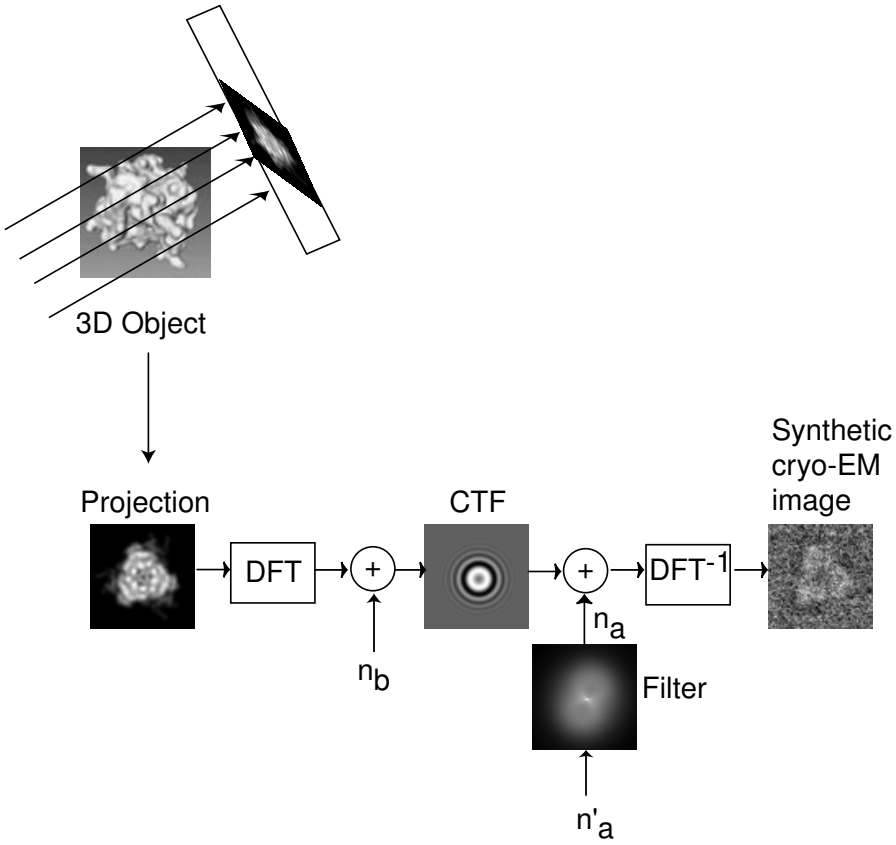


Figure 5.12: Procedure of synthesizing cryo-EM images.

our algorithm by the wavelet-projection matching.

Refinement

Starting from the initial misregistration of 4.50 ± 6.86 voxel, we performed the frequency-based assignment with adjusted standard deviations of the Gaussian functions, space-limiting, and cost-weighting, such that a good registration accuracy could be achieved. The standard deviation in each dimension for each Gaussian was 30.

The mean and the standard deviation of the misregistrations are shown in

Projections.

Projections after having applied the noise and the CTF.

Figure 5.13: Projections of the 3D model of the Hexon and the corresponding synthetic cryo-EM images. Courtesy of the Centro Nacional de Biotecnología, Madrid, Spain.

Table 5.1: Accuracy of the Hexon pose assignment by three standard angular refinement algorithms.

	Misregistration [voxel]	
	Mean	Standard deviation
Wavelet-projection matching	4.50	6.86
Space-projection matching	9.13	14.05
Radon angular assignment	35.00	11.59

Table 5.2. The frequency-based assignment was performed with a misregistration of 3.38 ± 7.64 voxel. The initial mean misregistration was reduced by about 25%. This is a significant accuracy improvement (with 95% confidence). The histograms of the initial and final misregistrations are shown in Figure 5.14.

Table 5.2: Accuracy of the Hexon pose assignment by the frequency-based registration for the initial wavelet-projection matching.

	Misregistration [voxel]	
	Mean	Standard deviation
Wavelet-projection matching	4.50	6.86
Frequency-based registration	3.38	7.64

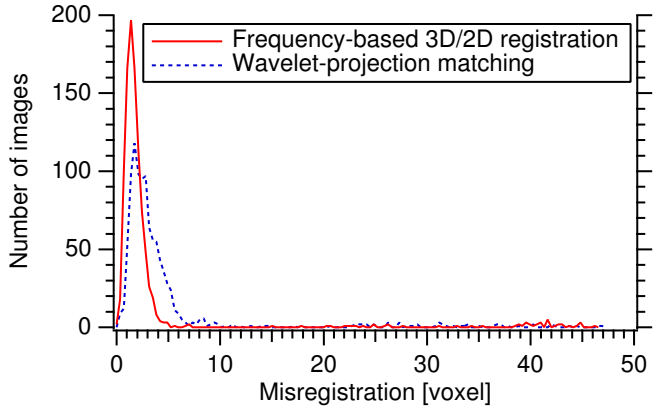


Figure 5.14: Histogram of the Hexon data misregistration. The frequency-based registration vs. the initial wavelet-projection matching.

5.4 Validation using data with an unknown ground-truth alignment

We now show the performance of our frequency-based algorithm in refining a 3D model of a GroEL chaperonin. The input for our algorithm was a set of experimental cryo-EM images, a 3D model of the GroEL, and an initial assignment for the particle poses obtained by using the wavelet-projection matching. In this case, there is no ground-truth alignment of the cryo-EM images to the 3D model. Besides, the input volume may be a completely wrong model of the protein. This means that we cannot evaluate objectively the accuracy of the assignment.

There have been many attempts in the literature on 3D EM to design a measure of resolution of a reconstructed protein model. The overview of the resolution measures is given in Section 5.4.1.

The cryo-EM images and the 3D model of the GroEL used for the experiment are described in Section 5.4.2. We evaluate the performance of our algorithm in Section 5.4.3. The results are given in terms of the Fourier Shell Correlation (FSC) which is the resolution measure that is most often used in practice.

5.4.1 Resolution

The theoretical resolution of the reconstruction that is expected from the data collection geometry is usually not achieved [26]. There are several reasons for this. One reason is that the structure of the particle is not defined to that level of resolution because of conformational variability. In stained preparations, for instance, the stain fluctuations limit the definition of small specimen features. Also, different electron optical effects (partial coherence, charging, specimen movement) limit the transfer of information from the specimen to the image. Errors in the angular assignment are additional limiting factors. The significant resolution is the resolution up to which object-related features are represented in the 3D reconstruction [26].

Measures of resolution

The resolution measures can be classified into two groups. Differential Phase Residual (DPR) [29] and Fourier Ring Correlation (FRC) [87] belong to the first group. Initially developed as measures of resolution for correlation-averaged images, they assess resolution in terms of the spectral range of consistency between the FTs of the averages computed for two data subsets. Similarly, 3D extensions of FRC and DPR assess resolution of the 3D reconstruction in terms of the spectral range of consistency between the FTs of the 3D reconstructions computed for two data subsets. The FSC is the 3D counterpart of the FRC [38]. The disadvantage of these measures is that their statistical properties are not well understood and that they are not true resolution measures but reproducibility measures.

The second group contains Q-factor [112] and Spectral Signal-to-Noise Ratio (SSNR) [72, 108, 109]. In the case of 2D data, they compare the FTs of all images participating in the average to the FT of the average. The statistical properties of the data are therefore directly captured. Q-factor is rarely used since it does not give results that can be expressed as a 1D function of frequency.

2D SSNR was introduced in [108]. It measures the SNR as a function of spatial frequency, and it provides a framework for relating FRC and DPR to each other in a meaningful way.

In [72], 3D SSNR was derived for a particular class of 3D reconstruction algorithms that use interpolation in the FT domain. An SSNR measure that can be used with any reconstruction algorithm has been proposed in [109]. This is a 2D approach to assess resolution of the 3D reconstruction. It is based on the comparison of the experimental projections to their predictions computed from the 3D reconstruction (estimate of the signal present in the experimental data). The article proposes to characterize the noise-reducing effect of the 3D reconstruction algorithm by introducing the noise reduction factor which is a function of spatial frequency, but also depends on the type of the reconstruction algorithm, the angular assignment, the type of symmetry, and the number of views. This factor cannot be determined analytically in general. The authors propose to estimate it in a two-step procedure. In the first step, they reconstruct a volume out of noise-only images with the same angular assignment as for the experimental images. In the second step, they get the estimate by computing the SSNR for the noise-only images. The final SSNR is the ratio between the SSNR for the experimental images and the SSNR for the noise-only images.

The 2D approaches have been criticized for not giving a fair assessment of 3D resolution since they do not consider all spatial directions. Since the study of 3D SSNR measures is still research in progress, 3D extensions of FRC and DPR are commonly used in practice.

3D DPR and FSC

The experimental images are randomly divided into two subsets of equal size. One set is commonly formed from even-numbered images, and another set from odd-numbered images of the data set. Two volumes are reconstructed, one for each subset. The resulting volumes are compared along shells in Fourier space. Let the 3D DFT of the two volumes be denoted by $\hat{f}_1$ and $\hat{f}_2$, and let $\mathbf{k}$ be the 3D frequency on the regular Fourier grid within the Nyquist range. If $\Delta\vartheta$ is the phase difference between the two transforms, then DPR is defined as

$$DPR(k, \Delta k) = \left| \frac{\sum_{[k,\Delta k]} (\Delta\vartheta(\mathbf{k}))^2 (|\hat{f}_1(\mathbf{k})| + |\hat{f}_2(\mathbf{k})|)}{\sum_{[k,\Delta k]} (|\hat{f}_1(\mathbf{k})| + |\hat{f}_2(\mathbf{k})|)} \right|^{1/2}, \tag{5.11}$$

whereas FSC is defined as

$$FSC(k, \Delta k) = \frac{\mathrm{Re}\left\{\sum_{[k,\Delta k]} \hat{f}_1(\mathbf{k})\hat{f}_2^*(\mathbf{k})\right\}}{\left|\sum_{[k,\Delta k]} |\hat{f}_1(\mathbf{k})|^2 \sum_{[k,\Delta k]} |\hat{f}_2(\mathbf{k})|^2\right|^{1/2}}. \tag{5.12}$$

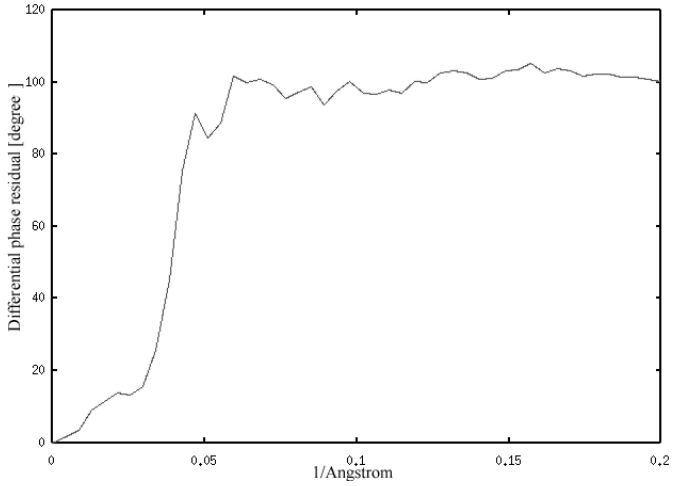
The sums are computed over shells defined by radii $k \pm \Delta k$, where $k = |\mathbf{k}|$. The two measures are plotted as functions of k. The two curves show the degree of consistency between the two volumes. The DPR curve starts with a value of 0 which indicates perfect agreement at zero frequency, and raises up with increasing frequency which indicates poor agreement for higher frequencies. In contrast, the FSC curve starts with a value of 1 which indicates perfect agreement at zero frequency, and falls gradually down which indicates poor agreement for higher frequencies. A typical behavior of the two measures can be seen in Figure 5.15. As a single DPR figure of merit, one usually selects the Fourier radius k for which $DPR(k, \Delta k) = 45°$. As a single FSC figure of merit, two cutoffs are usually considered. One is the radius k for which $FSC(k, \Delta k) = 0.5$, while the other is the radius for which the FSC curve intersects the $(3 - \sigma)$ curve [27]. The $(3 - \sigma)$ curve is obtained by multiplying the FSC curve for pure noise by 3.

The three figures have different values for the same data. The least optimistic is the 45°-DPR criterion, whereas the most optimistic is the $(3 - \sigma)$-FSC criterion. Unfortunately, there is no formal justification for using any of the three cutoffs. Besides, the quotation of a single number cannot do justice to the information conveyed by the whole curve [27]. The use of the $(3 - \sigma)$-FSC criterion was questioned in [27] using the relationship between the FSC and the SNR. The SNR drops quite rapidly with a decreasing FSC. An FSC of 0.5 corresponds to an SNR of 1. This means that a $(3 - \sigma)$-FSC cutoff corresponds to an even lower SNR, which makes this kind of cutoff questionable since the information cannot be reliably distinguished from noise at these frequencies.

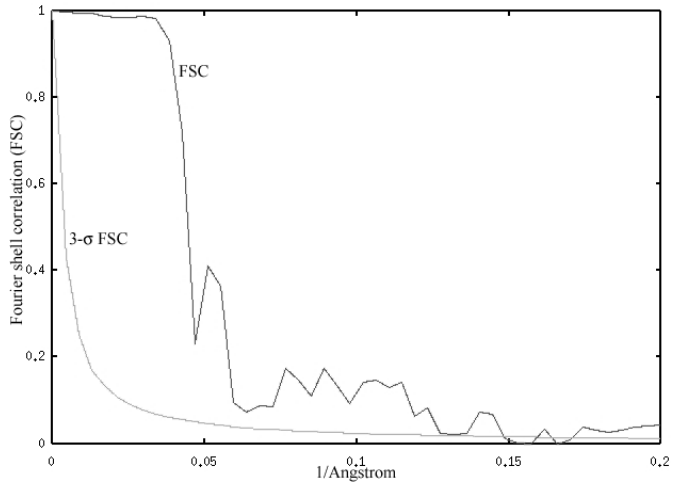
Here, we compute the FSC in the absence of a better measure of resolution.

5.4.2 Data

We used the cryo-EM images of a GroEL chaperonin from E. coli that have been acquired by S. De Carlo (Laboratoire d'analyse ultrastructurale, UNIL, Lausanne, Switzerland) as descibed in [22]. We used the 3D model of the GroEL that has been computed from these images by S. De Carlo as well. The 0.5-FSC of this volume of 14 $\mathring{A}$ was reported in [22].



Differential phase residual.



Fourier shell correlation.

Figure 5.15: Typical differential phase residual and Fourier shell correlation profiles.

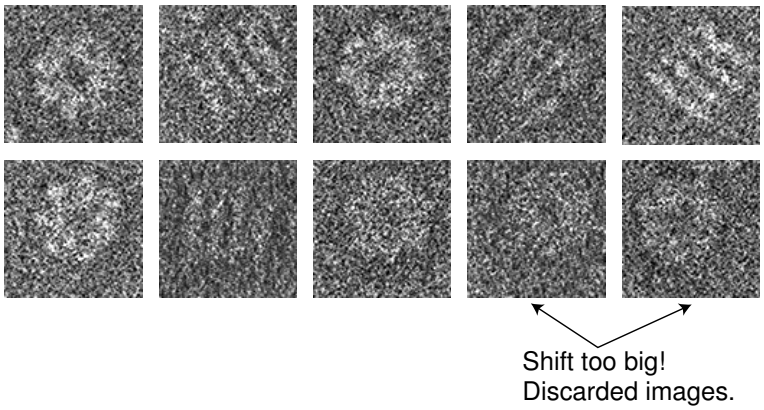


Figure 5.16: Cryo-EM images of GroEL. Courtesy of Laboratoire d'analyse ultrastructurale, UNIL, Lausanne, Switzerland.

GroEL chaperonin consists of two heptameric rings of identical subunits stacked back to back. In other words, it has a sevenfold rotational symmetry around one axis, and a twofold symmetry around one of the remaining axes. It assists protein folding in an ATP-regulated manner. Each subunit of GroEL contains an ATP binding site so that the entire GroEL has, therefore, 14 ATP binding sites.

We describe the procedure [22] used by S. De Carlo to obtain the cryo-EM images and the 3D model of the GroEL next.

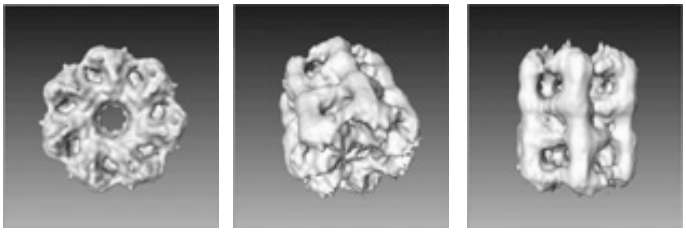
Cryo-EM images

GroEl particles were prepared for cryo-negative staining as described in [22]. Homemade holey carbon films were mounted on 200-mesh copper grids, one side of the grid was coated by a thin layer of Au/Pd, and the plastic film was dissolved with ethyl acetate. The grid with a drop of the sample applied onto the metal-coated side was staying in the staining solution for 30 s before it was mounted on the plunger, blotted, and vitrified in liquid ethane. The stain preparation is described elsewhere [22].

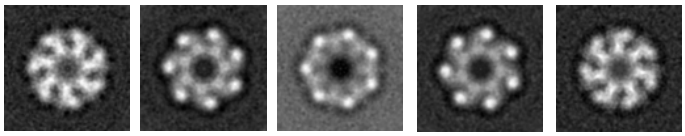
The micrographs were recorded at a temperature of $-180\,^{\circ}\mathrm{C}$ and at a magnification of 53,000 in a Philips CM100 transmission electron microscope (cathode source operated at 100 keV). A low-dose of $1{,}000\,\mathrm{e}^{-}/\mathrm{nm}^2$ was delivered to the sample in each exposure using the Philips beam-deflection unit.

The negatives were digitized using a Kodak RFS 3570 Scanner with a pixel size of 0.24 nm. The GroEL particles were picked using WEB, the graphic interface of the SPIDER package. The center of a window of size 96×96 pixel was manually placed onto the particle center, and the particle image of the same size was extracted from the micrograph. In this way, a set of about 2,500 particle images was created. Like described in [22], the images were contrast-normalized and no CTF correction was applied. Note that, although the CTF was not corrected, the particles stand out from the background (Figure 5.16), thanks to the technique of cryo-negative staining.

After discarding images that have excessive shift (for examples, see Figure 5.16), the total number of images for a further analysis was 2,322.



Three views of an isosurface.



A few slices.

Figure 5.17: 3D model of GroEL.

3D model

The images were multireference-aligned and classified using a multivariate statistical analysis, as implemented in the SPIDER package. The class averages were visualized in WEB. Since the class averages contained mainly top and side views of the GroEL chaperonin (this is also visible in Figure 5.16), the starting 3D model based on only these two views was computed in the Spider package using a back-projection reconstruction method.

The space-projection matching was applied to refine the starting model as explained in [22]. The volume used for the evaluation of our algorithm is the result of six refinement cycles. The 3D reconstruction is computed so that it takes into account the GroEL symmetries. A 3D visualization of this volume of size $96 \times 96 \times 96$ voxel is given in Figure 5.17.

5.4.3 Experiments

Our frequency-based algorithm used the 3D model from Section 5.4.2 and the initial pose assignment by the wavelet-projection matching. The initial assignment based on the same 3D model from Section 5.4.2 was performed by C. Ó. S. Sorzano.

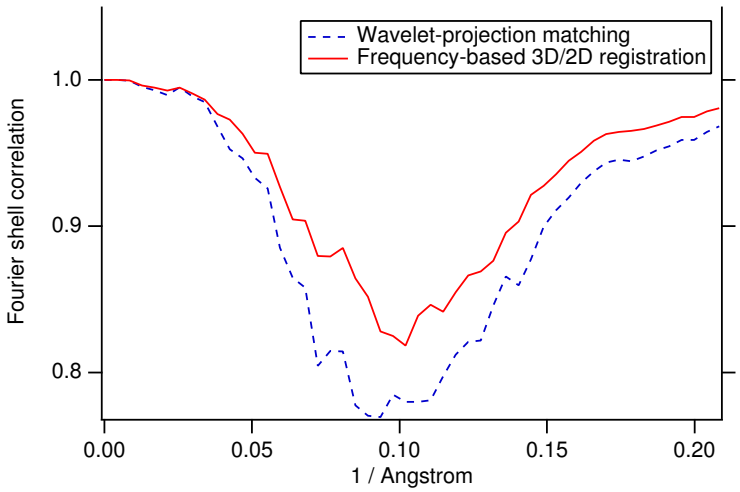


Figure 5.18: Fourier shell correlation for the GroEL assignment using the frequency-based method initialized by wavelet-projection matching.

Using the pose assignment by the wavelet-projection matching, we reconstructed two volumes from two subsets of the GroEl cryo-EM images. After having performed the frequency-based assignment, we reconstructed another two volumes from two subsets of the GroEl cryo-EM images. All four volume reconstructions were accomplished by C. Ó. S. Sorzano using ART with smooth spherically symmetric volume elements (Section 5.1.7).

Next, we computed the FSC curves for the two assignments, the initial one and the final one. The two FSC curves are shown in Figure 5.18. A striking feature is that the FSC in either of the two cases does not follow its "common" textbook behavior. It does not cross 0.5 at any frequency. Moreover, it increases at higher frequencies instead of falling down. An independent analysis has shown that it was possible to have two very consistent volumes at all frequencies using better-qualities reconstruction algorithms as the one used here. This is a good example for not trusting the FSC as the resolution measure but only as the consistency measure.

Therefore, the only conclusion that we can draw from Figure 5.18 is that the frequency-based algorithm improves the consistency of the initial volumes at all frequencies.

Chapter 6

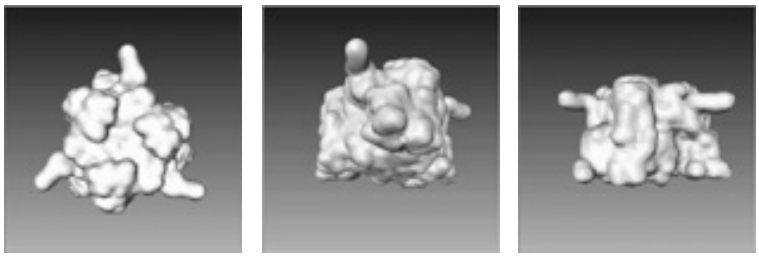
COMPARISON OF THE PERFORMANCE: SPACE VS. FREQUENCY

In Chapter 4, we showed the performance of the space-based algorithm using orthopedic data. In Chapter 5, we evaluated the performance of the frequency-based algorithm using electron-microscopy data. In this chapter, we compare the performance of the frequency-based algorithm with that of the space-based algorithm in the same electron-microscopy application. This application (Section 5.1) implies the parallel-beam geometry described in Chapter 3.

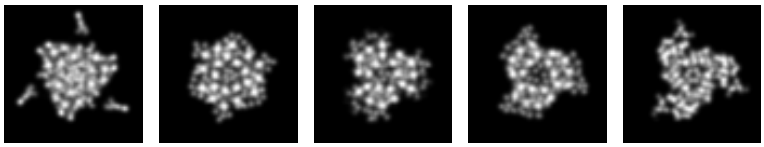
Let us recall that the frequency-based algorithm relies on the computation of the FTs of the volume projections using the CST approach and on a multiresolution weighting of the cost function (Chapter 3). The space-based algorithm is based on the computation of the volume projections by ray casting using the shearing method and on the spline multiresolution volume/image pyramids (Chapter 3).

We compare the quality of the assignment by the two algorithms using data with a known ground-truth alignment. For this purpose, we have repeated the procedure described in Section 5.3.2 to create synthetic cryo-EM images from a protein from the PDB. To measure the quality of the assignments, we employ the measure of misregistration given in Section 5.3.1.

The data are described in Section 6.1. In Section 6.2, we show two sets of experiments. The first one (Section 6.2.1) compares the robustness of the



Three views of an isosurface.



A few slices.

Figure 6.1: 3D model of the Bacteriorhodopsin from the PDB.

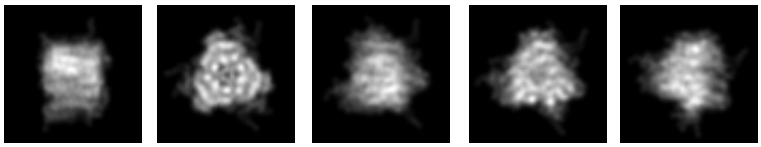
two algorithms to a random value for the initial assignment. The other set of experiments (Section 6.2.2) compares the ability of the two algorithms to refine the initial assignment obtained using the wavelet-projection matching. We compare the speed of the two approaches in Section 6.2.3. A discussion is found in Section 6.3.

6.1 Data

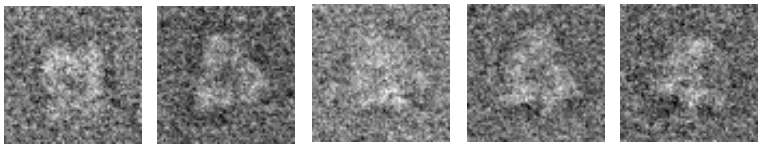
6.1.1 Volume

We used the Bacteriorhodopsin volume from the PDB (accession code: 1BRD [40]). The size of the volume was $64 \times 64 \times 64$ voxel and its sampling rate was 3.27 Å (Figure 6.1).

The Bacteriorhodopsin found in Halobacterium halobium is a membrane protein that functions as a light-driven pump of protons across the cell membrane. The proton gradient is used by a second membrane protein called ATP synthese to generate metabolic energy in the form of ATP which is used by the cell to drive many vital processes. When crystallized, the Bacteriorhodopsin forms hexagonal 2D crystals. They are sufficiently well ordered for determin-



Projections.



Projections after having applied the noise and the CTF.

Figure 6.2: Projections of the 3D model of the Bacteriorhodopsin and the corresponding synthetic cryo-EM images. Courtesy of the Centro Nacional de Biotecnología, Madrid, Spain.

ing a 3D structure at near-atomic resolution using high-resolution cryoelectron-microscopy [40]. This protein has a threefold rotational symmetry.

6.1.2 Images

The synthetic cryo-EM images of the Bacteriorhodopsin were synthesized by C. Ó. S. Sorzano (Centro Nacional de Biotecnología, Madrid, Spain) in the two-step procedure described in Section 5.3.2.

A few examples of these random projections and the corresponding synthesized cryo-EM images can be seen in Figure 6.2.

6.2 Experiments

We computed three-level dyadic volume and image pyramids for the space-based registration. As far as the frequency-based method is concerned, we have adjusted the standard deviations of the Gaussian functions, space-limiting and cost-weighting as described in Section 5.3.3.

6.2.1 Robustness

We designed four experiments to test the robustness of the two algorithms on the initial assignment. In the first experiment, we initialized the parameters by their ground-truth values. In the second experiment, we initialized the angles by their ground-truth values modified by 5° or -5° with probability 0.5. The initial values for the translation parameters were equal to their ground-truth values. In the third and fourth experiments, the angular perturbation was $\pm 10^\circ$ and $\pm 15^\circ$, respectively. The initial misregistration is reported in Table 6.1 by its mean value and its standard deviation.

Table 6.1: Initial Bacteriorhodopsin data misregistration for the robustness comparisons.

Initial assignment	Misregistration [voxel]	
	Mean	Standard deviation
Ground truth	0	0
Ground truth $\pm 5^\circ$	3.55	0.89
Ground truth $\pm 10^\circ$	7.10	1.76
Ground truth $\pm 15^\circ$	10.55	2.61

Table 6.2: Robustness of the space-based algorithm. The Bacteriorhodopsin data misregistration.

Space-Based	Misregistration [voxel]	
	Mean	Standard deviation
Ground truth	1.25	0.58
Ground truth $\pm 5^\circ$	1.89	0.77
Ground truth $\pm 10^\circ$	4.30	2.22
Ground truth $\pm 15^\circ$	7.86	3.93

The space-based method assigned the poses with a misregistration whose mean and standard deviation are shown in Table 6.2. The residual data misregistration by the frequency-based method is shown in Table 6.3.

Starting from the ground-truth assignment, the two algorithms performed in a similar way (with a confidence of 95%). Both algorithms converged to

Table 6.3: Robustness of the frequency-based algorithm. The Bacteriorhodopsin data misregistration.

Frequency-based	Misregistration [voxel]	
	Mean	Standard deviation
Ground truth	1.30	0.77
Ground truth $\pm 5°$	1.61	0.85
Ground truth $\pm 10°$	1.83	1.31
Ground truth $\pm 15°$	2.53	2.91

an assignment that was not the ground-truth one. This can be attributed to the presence of noise and to the CTF, which make a synthetic cryo-EM image differ from the corresponding theoretical projection. But when the angular perturbation was substantial, both algorithms were successful in reducing the initial misregistration. For the angular perturbation of $\pm 5°$, the frequency-based method achieved a registration that was slightly more accurate (with a confidence of 95%) than the one achieved by the space-based method. In the two remaining cases, $\pm 10°$ and $\pm 15°$, the frequency-based method was much more accurate (with a confidence of 95%) than the space-based method.

6.2.2 Refinement of a standard-method assignment

Initialization

In this section, we initialize our frequency- and space-based algorithms using the wavelet-projection matching since it showed the best performance when compared with two other standard refinement algorithms.

For completeness purposes, we show here the results of this comparison for the Bacteriorhodopsin data. This comparison was done by C. Ó. S. Sorzano [93]. The mean and the standard deviation of the Bacteriorhodopsin data misregistration after having run the three standard algorithms are summarized in Table 6.4. The values of the mean and standard deviation of the misregistration are computed in two ways. One uses the assignments for all 1,000 synthetic images while the other takes into account only the images that were misregistered with respect to the volume for not more than 10 voxel (that we consider as reasonably assigned images). A misregistration of 10 voxel corresponds to an angular perturbation of about 15 degree around the ground-truth angular

assignment. The number of images taken into account is also reported in Table 6.4.

The proportions of the reasonably assigned images were [98.7%,99.8%], [93.2%,96.1%], and [59.3%,65.4%] for the wavelet-projection matching, the space-projection matching, and the Radon angular assignment, respectively, which means that the wavelet-projection matching was the most robust (with 95% confidence) among the three methods. According to the mean value of the misregistration, the wavelet-projection matching appears to have a lesser accuracy than the space-projection matching. However, when statistical significance is taken into consideration (with 95% confidence), these two methods are similar. Both also perform significantly better than the Radon angular assignment.

Table 6.4: Accuracy and robustness of the Bacteriorhodopsin pose assignment by three standard refinement algorithms.

	Misregistration [voxel]		Number of images
	Mean	Standard deviation	
Wavelet-projection matching	3.18	1.85	1,000
Space-projection matching	2.99	3.68	1,000
Radon angular assignment	16.55	18.80	1,000

All synthetic images taken into account.

	Misregistration [voxel]		Number of images
	Mean	Standard deviation	
Wavelet-projection matching	3.12	1.65	994
Space-projection matching	2.25	1.58	948
Radon angular assignment	3.13	2.09	624

Synthetic images with a misregistration of
less than 10 voxel taken into account.

Refinement

The goal of our algorithms was to reduce the initial data misregistration of 3.18 ± 1.85 voxel. The mean values and the standard deviations of the initial and final data misregistrations for both algorithms are shown in Table 6.5.

Both algorithms were successful in reducing (with 95% confidence) the ini-

Table 6.5: Accuracy of the frequency- and space-based algorithms for the initial wavelet-projection matching (Bacteriorhodopsin).

	Misregistration [voxel]		Number of images
	Mean	Standard deviation	
Wavelet-projection matching	3.18	1.85	1,000
Space-based registration	2.23	1.54	1,000
Frequency-based registration	1.77	1.11	1,000

tial misregistration. The frequency-based method registered the volume to the images with an accuracy of 1.77 ± 1.11 voxel which was slightly better (with a confidence of 95%) than the accuracy of the space-based method (2.23 ± 1.54 voxel). The space-based method reduced the initial mean misregistration by about 30%, while the frequency-based method reduced it by 44%.

We conclude that the frequency-based volume-to-image registration algorithm was more accurate than the space-based method. These results are consistent with the results obtained from experiments on the robustness of the two algorithms which showed that the frequency-based algorithm was more robust than the space-based method.

Robustness to the angular step of the library

We explored the effect of a coarser library of projections, used by the wavelet-projection matching, on the performance of the frequency-based algorithm. We did not consider the space-based method since it seemed slightly less robust than the frequency-based method. We created a library of 150 projections of the Bacteriorhodopsin with an angular step of 10°. Then, we performed the assignment using the wavelet-projection matching. We run the frequency-based algorithm with the volume and the images initially misregistered by 3.85 ± 3.68 voxel. The registration was accomplished with an accuracy of 1.81 ± 1.44 voxel.

We can notice that the initial misregistration was significantly higher (with 95% confidence) than the misregistration used to initialize the registration based on the 5°-library. In both cases, we significantly reduce (with 95% confidence) the initial misregistration (Table 6.6). The histograms of the misregistration (with a bin width of 0.35 voxel) in the two cases are shown in Figure 6.3.

Table 6.6: Influence of a 5°-library and a 10°-library of the volume projections, used for the initial wavelet-projection matching, on the performance of the frequency-based algorithm (Bacteriorhodopsin).

5°	Misregistration [voxel]	
	Mean	Standard deviation
Wavelet-projection matching	3.18	1.85
Frequency-based registration	1.77	1.11

5°-Library.

10°	Misregistration [voxel]	
	Mean	Standard deviation
Wavelet-projection matching	3.85	3.68
Frequency-based registration	1.81	1.44

10°-Library.

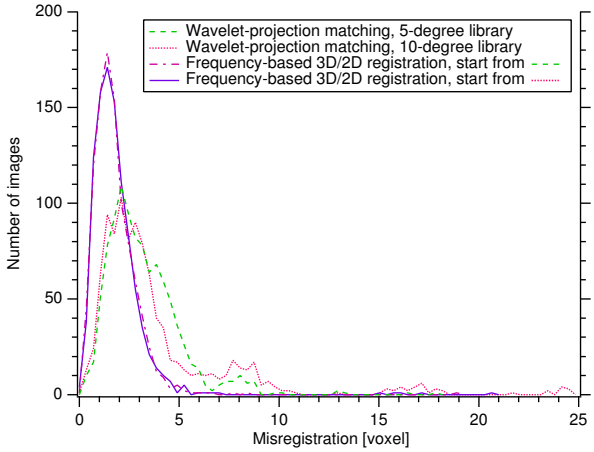


Figure 6.3: Histogram of the Bacteriorhodopsin data misregistration. Influence of a 5°-library and a 10°-library of the volume projections, used for the initial wavelet-projection matching, on the performance of the frequency-based algorithm.

6.2.3 Speed

For the image size 64×64 pixel and the volume size $64 \times 64 \times 64$ voxel, the frequency-based registration takes about 12 sec while the space-domain registration takes about 17 sec per image on a Power Mac G4, 733MHz. The computation of the volume and image pyramids takes approximately 1.8 sec and 0.2 sec, respectively.

6.3 Discussion

We compared the space-based algorithm with the frequency-based algorithm in a fully controlled simulation environment where the ground-truth solution was known *a priori*. We synthesized a set of images at known poses using a 3D model of a protein from the PDB. The synthesized images were very similar to the data commonly encountered in 3D electron cryo-microscopy. Since they have a nearly limited extent in the space domain, their accurate interpolation in the Fourier domain can be achieved using sinc-like basis functions. Our algorithms are based on cubic-spline interpolation. The cardinal cubic spline is quite similar to the sinc function as we have seen in Section 2.2. This is an explanation for the gain in accuracy achieved by the frequency-based registration method over the space-based method.

Chapter 7

CONCLUSIONS

7.1 Summary

In this thesis, we have described two novel algorithms for the registration of a volumetric image to a set of given 2D projections. The first algorithm performs the registration in the space domain, while the second does it in the frequency domain. They rely on the iterative minimization of the least-squares difference between the given projections and simulated projections of the volume in the corresponding domain. The main novelty of our algorithms is the use of cubic splines for performing interpolation, either in the space domain or in the frequency domain. This allows for a well-defined gradient of the dissimilarity measure, which is a necessary condition for efficient and accurate registration.

In a first step, we have developed our space-based and frequency-based registration strategies in the context of the axonometric projection. We have focused on methods for a fast simulation of volume projections and on multiresolution strategies for improvement of the robustness. We have used ray casting to simulate the volume projections for the space-based registration. To speed-up the computations, we proposed a novel one-step approach for fast ray casting that requires mere 2D interpolation instead of the more costly 3D interpolation. To improve the robustness of the space-based approach, we proposed a multiresolution optimization strategy where spline-based data pyramids are processed in coarse-to-fine fashion, which improves speed as well. In the context of axonometric projection, we have also discussed the use of the central-slice theorem to determine a signature of the simulated projections; this is at the core of

our frequency-based registration algorithm. Then, Parseval's theorem allows for a fast computation of the dissimilarity measure. To improve robustness, we proposed to emulate the coarse-to-fine strategy by applying weights in the frequency domain.

We have also adapted our space-based algorithm to the perspective projection model, sometimes called a conical-beam projection. This allows us to apply it to the registration of a CT volume to a set of C-arm images for computer-assisted orthopedic surgery. We observe that the registration accuracy achieved in this case is consistent with the current standards. When registering the CT to pairs of real C-arm images, we obtain an accuracy of about 1.7 ± 0.5 mm, which is subvoxel with respect to the inter-slice CT thickness of 2.5 mm. We have shown that the registration accuracy could be improved by increasing the number of C-arm images. The use of three C-arm images results in a registration accuracy of 1.4 ± 0.2 mm.

Then, we have applied our frequency-based registration algorithm to the 3D reconstruction of a protein from its 2D electron-microscopy images, which is an application that is perfectly suited to the axonometric projection. We have shown that our algorithm could be used to obtain a refined 3D model of a protein with respect to currently available algorithms. In this case, we achieve a better registration accuracy because we deal with a continuous space of geometric parameters, while the standard methods achieve only limited accuracy because they deal with quantized parameters. In the experiments on synthetic bacteriorhodopsin data, our algorithm reduces the initial misregistration achieved by wavelet-projection matching by about 44%. In case of less cooperative hexon data, our algorithm refines the solution by 25%. We have also performed experiments with real protein data. They show that our algorithm improves their consistency at all frequencies.

Last, we have compared the performance of the two algorithms in the context of electron microscopy, in a fully controlled simulation environment where the ground-truth solution was known *a priori*. The frequency-based registration was slightly more accurate than the space-based one. This result can be attributed to the suitability of interpolation in the frequency domain to data that are strictly limited in space.

7.2 Possible extensions

Space-based registration

• Explore the use of other dissimilarity measures. For instance, the mutual information has been shown to suit well the task of registering multimodal data.
• Investigate the use of weights that are related to the in-volume ray length. Note that the in-volume ray length depends on the volume pose, which means that the weight given to each pixel will change in each iteration.
• Extend the rigid-body transformation to elastic transformations.

Frequency-based registration

• Substitute the Fourier domain by some other domain for which the central-slice theorem holds, for instance the Hartley domain, which has the advantage of being real-valued.
• Consider the perspective projection, which requires the development of a formulation of the central-slice theorem that is tuned to the perspective geometry.

Applications

Computer-assisted orthopedic surgery

• Validate with other data sets (simple-geometry phantom data, missing data in one of the two modalities, patient data).

3D electron-microscopy

• Validate with other real-data sets.
• Investigate the use of the spectral signal-to-noise ratio as a measure of resolution of a 3D reconstruction instead of the Fourier-shell correlation.

Other applications

• Interventional neuroradiology (registration of 3D x-ray angiography data or 3D magnetic resonance angiography data to 2D digital subtracted angiography images of vessels in the brain or kidney)
• Breast pathology assessment (registration of 3D magnetic resonance data to 2D x-ray mammography images of the breast).

Bibliography

[1] A. Aldroubi, M. Unser, and M. Eden, *Cardinal spline filters: Stability and convergence to the ideal sinc interpolator*, Signal Processing, 28 (1992), pp. 127–138.

[2] E. Bainville, I. Bricault, P. Cinquin, and S. Lavallée, *Concepts and methods of registration for computer-integrated surgery*, in Computer Assisted Orthopedic Surgery (CAOS), L. P. Nolte and R. Ganz, eds., Hogrefe & Huber Publishers, Seattle-Toronto-Bern-Göttingen, 1999, pp. 15–34.

[3] M. Bajura and U. Neumann, *Dynamic registration correction in video-based augmented reality systems*, IEEE Computer Graphics and Applications, 15 (1995), pp. 52–60.

[4] R. Bansal, L. H. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, and J. S. Duncan, *Entropy-based, multiple-portal-to-3DCT registration for prostate radiotherapy using iteratively estimated segmentation*, in Proc. MICCAI'99, C. Taylor and A. Colchester, eds., 1999, pp. 567–578.

[5] S. Basu and Y. Bresler, *Feasibility of tomography with unknown view angles*, IEEE Trans. Image Processing, 9 (2000), pp. 1107–1122.

[6] ——, *Uniqueness of tomography with unknown view angles*, IEEE Trans. Image Processing, 9 (2000), pp. 1094–1106.

[7] U. Berlemann, M. Slomczykowski, F. Langlotz, R. Hofstetter, M. Sati, H. Visarius, and L.-P. Nolte, *Computerassistenz in der orthopädischen chirurgie-Grundlegende konzepte und weiterentwicklungen*

durch integration der fluoroskopie, Sportorthopädie-Sporttraumatologie, 15 (1999), pp. 155–159.

[8] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999.

[9] E. M. Bispo and R. B. Fisher, *Free-form surface matching for surface inspection*, in The Mathematics of Surfaces VI, G. Mullineux, ed., Oxford Clarendon Press, 1996, pp. 119–136.

[10] T. Blu, P. Thévenaz, and M. Unser, *MOMS: Maximal-order interpolation of minimal support*, IEEE Trans. Image Processing, 10 (2001), pp. 1069–1080.

[11] T. Blu and M. Unser, *Approximation error for quasi-interpolators and (multi-) wavelet expansions*, Appl. Comput. Harmonic Anal., 6 (1999), pp. 219–251.

[12] ———, *Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors*, IEEE Trans. Signal Processing, 47 (1999), pp. 2783–2795.

[13] ———, *Quantitative Fourier analysis of approximation techniques: Part II—Wavelets*, IEEE Trans. Signal Processing, 47 (1999), pp. 2796–2806.

[14] P. Brigger, F. Müller, K. Illgner, and M. Unser, *Centered pyramids*, IEEE Trans. Image Processing, 8 (1999), pp. 1254–1264.

[15] L. G. Brown, *A survey of image registration techniques*, ACM Computing Surveys, 24 (1992), pp. 325–376.

[16] L. M. G. Brown and T. E. Boult, *Registration of planar film radiographs with computed tomography*, in IEEE Proc. MMBIA'96, 1996, pp. 42–51.

[17] P. E. Carlsson, P. R. Edholm, and P. E. Danielsson, *The synthesis of a new X-ray picture identical in projection to a previous picture using 3D Fourier techniques*, Phys. Med. Biol., 39 (1994), pp. 597–608.

[18] C. R. Crawford, J. G. Colsher, N. J. Pelc, and A. H. R. Lonn, *High speed reprojection and its applications*, in Proc. SPIE. Medical Imaging II, vol. 914, 1988, pp. 311–318.

[19] C. M. Cyr, A. F. Kamal, T. B. Sebastian, and B. B. Kimia, *2D-3D registration based on shape matching*, in Proc. IEEE Workshop MMBIA-2000, 2000, pp. 198–203.

[20] A. Czopf, C. Brack, M. Roth, and A. Schweikard, *3D-2D registration of curved objects*, Periodica Polytechnica Ser. El. Eng., 43 (1999), pp. 19–41.

[21] J. Dengler, *A multi-resolution approach to the 3D reconstruction from an electron microscope tilt series solving the alignment problem without gold particles*, Ultramicroscopy, 30 (1989), pp. 337–348.

[22] S. De Carlo, C. El-Bez, C. Alvarez-Rúa, J. Borge, and J. Dubochet, *Cryo-negative staining reduces electron-beam sensitivity of vitrified biological particles*, J. Struct. Biol., 138 (2002), pp. 216–226.

[23] S. Dunne, S. Napel, and B. Rutt, *Fast reprojection of volume data*, in Proc. First Conf. Visualiz. Biomed. Comput., 1990, pp. 11–18.

[24] P. J. Edwards, D. J. Hawkes, G. P. Penney, and M. J. Clarkson, *Guiding therapeutic procedures*, in Medical Image Registration, J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, eds., CRC Press LLC, Boca Raton, Florida, 2001, pp. 253–278.

[25] J. Feldmar, N. Ayache, and F. Betting, *3D-2D projective registration of free-form curves and surfaces*, Comput. Vision Image Understanding, 65 (1997), pp. 403–424.

[26] J. Frank, *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*, Academic Press, 1996.

[27] ——, *Single-particle imaging of macromolecules by cryo-electron microscopy*, Annu. Rev. Biophys. Biomol. Struct. 2002, 31 (2002), pp. 303–319.

[28] J. Frank, M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith, *SPIDER and WEB: Processing and visualization of images in 3-D electron microscopy and related fields*, J. Struct. Biol., 116 (1996), pp. 190–199.

[29] J. Frank, A. Verschoor, and M. Boublik, *Computer averaging of electron micrographs of 40S ribosomal subunits*, Science, 214 (1981), pp. 1353–1355.

[30] M. S. Gelfand and A. B. Goncharov, *Spatial rotational alignment of identical particles given their projections: Theory and practice*, Transl. Math. Monographs, 81 (1990), pp. 97–122.

[31] K. G. A. Gilhuijs, P. J. H. van de Ven, and M. van Herk, *Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data*, Med. Phys., 23 (1996), pp. 389–399.

[32] A. B. Goncharov, *Three-dimensional reconstruction of arbitrarily arranged identical particles given their projections*, Transl. Math. Monographs, 81 (1990), pp. 67–95.

[33] R. Grosso and T. Ertl, *Biorthogonal wavelet filters for frequency domain volume rendering*, in Visualization in Scientific Computing '95. Proceedings of the Eurographics Workshop, R. Scateni, J. van Wijk, and P. Zanarini, eds., 1995, pp. 81–95.

[34] A. Guéziec, P. Kazanzides, B. Williamson, and R. H. Taylor, *Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot*, IEEE Trans. Med. Imag., 17 (1998), pp. 715–728.

[35] A. Hamadeh and P. Cinquin, *Kinematic study of lumbar spine using functional radiographies and 3D/2D registration*, in Proc. CVRMed-MRCAS'97, J. Troccaz, E. Grimson, and R. Mosges, eds., 1997, pp. 109–118.

[36] A. Hamadeh, P. Sautot, S. Lavallée, and P. Cinquin, *Towards automatic registration between CT and X-ray images: Cooperation between 3D/2D registration and 2D edge detection*, in Proc. Med. Robot. Comput.-Assisted Surgery (MRCAS'95), 1995, pp. 39–46.

[37] G. Harauz and F. P. Ottensmeyer, *Direct three-dimensional reconstruction for macromolecular complexes from electron micrographs*, Ultramicroscopy, 12 (1984), pp. 309–319.

[38] G. Harauz and M. van Heel, *Exact filters for general geometry three-dimensional reconstruction*, Optik, 73 (1986), pp. 146–156.

[39] D. J. Hawkes, *Registration methodology: Introduction*, in Medical Image Registration, J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, eds., CRC Press LLC, Boca Raton, Florida, 2001, pp. 11–38.

[40] R. Henderson, J. M. Baldwin, T. A. Ceska, F. Zemlin, E. Beckmann, and K. H. Downing, *Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy*, J. Mol. Biol., 213 (1990), pp. 899–929.

[41] G. T. Herman and A. Kuba, eds., *Discrete Tomography: Foundations, Algorithms, and Applications*, Boston Birkhäuser, 1999.

[42] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis—Foundations*, Oxford University Press Inc., New York, 1996.

[43] R. Hofstetter, M. Slomczykowski, M. Sati, and L.-P. Nolte, *Fluoroscopy as an imaging means for computer-assisted surgical navigation*, Comp. Aided Surg., 4 (1999), pp. 65–76.

[44] A. Imiya, *The identification problem and image reconstruction of the exponential X-ray transform*, Trans. Inst. Elec. Inform. Comm. Eng. A., J75-A (1992), pp. 831–839.

[45] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Inc., 1989.

[46] S. Jonić, P. Thévenaz, and M. Unser, *Multiresolution spline-based 3D/2D registration of CT volume and C-arm images for computer-assisted surgery*, in Proc. SPIE. Medical Imaging 2001: Image Processing, M. Sonka and K. M. Hanson, eds., vol. 4322, 2001, pp. 1101–1109.

[47] ——, *Multiresolution-based registration of a volume to a set of its projections*, in Proc. SPIE. Medical Imaging 2003: Image Processing, M. Sonka and J. M. Fitzpatrick, eds., vol. 5032, 2003, pp. 1049–1052.

[48] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.

[49] C. T. Kelley, *Iterative Methods for Optimization*, Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999.

[50] W. Kühlbrandt and K. A. Williams, *Analysis of macromolecular structure and dynamics by electron cryo-microscopy*, Current Opinion in Chemical Biology, 3 (1999), pp. 537–543.

[51] D. LaRose, J. Bayouth, and T. Kanade, *Transgraph: Interactive intensity-based 2D/3D registration of X-ray and CT data*, in Proc. SPIE. Medical Imaging 2000: Image Processing, K. M. Hanson, ed., vol. 3979, 2000, pp. 385–396.

[52] P. D. Lauren and N. Nandhakumar, *Estimating the viewing parameters of random, noisy projections of asymmetric objects for tomographic reconstruction*, IEEE Trans. Pattern Anal. Machine Intell., 19 (1997), pp. 417–430.

[53] S. Lavallée and R. Szeliski, *Recovering the position and orientation of free-form objects from image contours using 3D distance maps*, IEEE Trans. Pattern Anal. Machine Intell., 17 (1995), pp. 378–390.

[54] L. Lemieux, R. Jagoe, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas, *A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs*, Med. Phys., 21 (1994), pp. 1749–1760.

[55] M. Levoy, *Volume rendering using the Fourier projection-slice theorem*, in Proc. Graphics Interface '92, 1992, pp. 61–69.

[56] R. M. Lewitt, *Multidimensional digital image representations using generalized Kaiser-Bessel window functions*, J. Opt. Soc. Am. A, 7 (1990), pp. 1834–1846.

[57] ——, *Alternatives to voxels for image representation in iterative reconstruction algorithms*, Phys. Med. Biol., 37 (1992), pp. 705–716.

[58] T. Malzbender, *Fourier volume rendering*, ACM Trans. Graph., 12 (1993), pp. 233–250.

[59] R. Marabini, G. T. Herman, and J. M. Carazo, *3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs)*, Ultramicroscopy, 72 (1998), pp. 53–65.

[60] R. Marabini, I. M. Masegosa, S. M. M. C. SanMartín, J. J. Fernández, L. G. de la Fraga, C. Vaquerizo, and J. M. Carazo, *Xmipp: An image processing package for electron microscopy*, J. Struct. Biol., 116 (1996), pp. 237–240.

[61] S. Matej and R. M. Lewitt, *Efficient 3D grids for image reconstruction using spherically-symmetric volume elements*, IEEE Trans. Nucl. Sci., 42 (1995), pp. 1361–1370.

[62] ——, *Practical considerations for 3-D image reconstruction using spherically symmetric volume elements*, IEEE Trans. Med. Imaging, 15 (1996), pp. 68–78.

[63] C. Meilhac and C. Nastar, *Robust fitting of 3D CAD models to video streams*, in Proc. ICIAP'97, A. Del-Bimbo, ed., vol. 1, 1997, pp. 661–668.

[64] S. A. Meyer and P. D. Wolf, *Registration of three-dimensional cardiac catheter models to single-plane fluoroscopic images*, IEEE Trans. Biomed. Eng., 46 (1999), pp. 1471–1479.

[65] D. C. Munson, J. D. O'Brien, and W. K. Jenkins, *A tomographic formulation of spotlight-mode synthetic aperture radar*, Proc. IEEE, 71 (1983), pp. 917–925.

[66] H. C. Murrell and D. Carson, *Image reconstruction via the Hartley transform*, South African Computer Journal, 1 (1990), pp. 36–42.

[67] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons Ltd. and B G Teubner, Stuttgart, 1986.

[68] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*, Philadelphia, PA: Society for Industrial and Applied Mathematics, 2001.

[69] L.-P. Nolte, M. A. Slomczykowski, U. Berlemann, M. J. Strauss, R. Hofstetter, D. Schlenzka, T. Laine, and T. Lund, *A new approach to computer-aided spine surgery: Fluoroscopy-based surgical navigation*, Eur. Spine J., 9 (2000), pp. S78–S88.

[70] S. J. Norton, *Tomographic reconstruction of 2-D vector fields: Application to flow imaging*, Geophys. J., 97 (1989), pp. 161–168.

[71] P. Penczek, M. Radermacher, and J. Frank, *Three-dimensional reconstruction of single particles embedded in ice*, Ultramicroscopy, 40 (1992), pp. 33–53.

[72] P. A. Penczek, *Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms*, J. Struct. Biol., 138 (2002), pp. 34–46.

[73] P. A. Penczek, R. A. Grassucci, and J. Frank, *The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles*, Ultramicroscopy, 53 (1994), pp. 251–270.

[74] G. P. Penney, P. G. Batchelor, D. L. G. Hill, D. J. Hawkes, and J. Weese, *Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images*, Med. Phys., 28 (2001), pp. 1024–1032.

[75] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, *A comparison of similarity measures for use in 2-D-3-D medical image registration*, IEEE Trans. Med. Imag., 17 (1998), pp. 586–595.

[76] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes*, The Art of Scientific Computing, Cambridge University Press, 1988.

[77] S. W. Provencher and R. H. Vogel, *Three-dimensional reconstruction from electron micrographs of disordered specimens. I. Method*, Ultramicroscopy, 25 (1988), pp. 209–222.

[78] M. Radermacher, *Weighted back-projection methods*, in Electron Tomography: Three-Dimensional Imaging With the Transmission Electron Microscope, J. Frank, ed., Plenum Press, New York, 1992, pp. 91–115.

[79] ——, *Three-dimensional reconstruction from random projections: Orientational alignment via Radon transforms*, Ultramicroscopy, 53 (1994), pp. 121–136.

[80] J. W. Rector and J. K. Washbourne, *Characterization of resolution and uniqueness in crosswell direct-arrival traveltime tomography using the Fourier projection slice theorem*, Geophys., 59 (1994), pp. 1642–1649.

[81] T. Rohlfing, D. B. Russakoff, M. J. Murphy, and C. R. Maurer, *An intensity-based registration algorithm for probabilistic images and its application for 2-D to 3-D image registration*, in Proc. SPIE. Medical Imaging 2002: Image Processing, M. Sonka and J. M. Fitzpatrick, eds., vol. 4684, 2002, pp. 581–591.

[82] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, *Nonrigid registration using free-form deformations: Application to breast MR images*, IEEE Trans. Med. Imag., 18 (1999), pp. 712–721.

[83] D. B. Russakoff, T. Rohlfing, D. Rueckert, R. Shahidi, D. Kim, and C. R. Maurer, *Fast calculation of digitally reconstructed radiographs using light fields*, in Proc. SPIE. Medical Imaging 2003: Image Processing, M. Sonka and J. M. Fitzpatrick, eds., vol. 5032, 2003, pp. 684–695.

[84] J. J. Rux and R. M. Burnett, *Type-specific epitope locations revealed by X-ray crystallographic study of adenovirus type 5 hexon*, Molecular Therapy, 1 (2000), pp. 18–30.

[85] D. B. Salzman, *A method of general moments for orienting 2D projections of unknown 3D objects*, Comput. Vision Graphics Image Process, 50 (1990), pp. 129–156.

[86] D. Sarrut and S. Clippe, *Geometrical transformation approximation for 2D/3D intensity-based registration of portal images and CT scan*, in Proc. MICCAI 2001, W. Niessen and M. Viergever, eds., 2001, pp. 532–540.

[87] W. O. Saxton and W. Baumeister, *The correlation averaging of a regularly arranged bacterial envelope protein*, J. Microsc., 127 (1982), pp. 127–138.

[88] I. J. Schoenberg, *Contributions to the problem of approximation of equidistant data by analytic functions*, Quart. Appl. Math., 4 (1946), pp. 45–99,112–141.

[89] G. W. Sherouse, K. Novins, and E. Chaney, *Computation of digitally reconstructed radiographs for use in radiotherapy treatment design*, Int. J. Radiation Oncology Biol. Phys., 18 (1990), pp. 651–658.

[90] R. Singh, R. M. Voyles, D. Littau, and N. P. Papanikolopoulos, *Shape morphing-based control of robotic visual servoing*, Autonomous Robots, 10 (2001), pp. 317–338.

[91] L. M. Sirois, D. H. Hristov, and B. G. Fallone, *Three-dimensional anatomy setup verification by correlation of orthogonal portal images and digitally reconstructed radiographs*, Med. Phys., 26 (1999), pp. 2422–2428.

[92] C. O. S. Sorzano, *Algoritmos Iterativos de Tomografía Tridimensional en Microscopía Electrónica de Transmisión*, Univ. Politécnica de Madrid, 2002. (Ph.D. thesis).

[93] C. O. S. Sorzano, S. Jonić, C. El-Bez, J. M. Carazo, S. De Carlo, P. Thévenaz, and M. Unser, *A multiresolution approach to pose assignment in 3-D electron microscopy of single particles*, J. Struct. Biol. (submitted).

[94] N. Sugano, *Computer-assisted orthopedic surgery*, J. Orthop. Sci., 8 (2003), pp. 442–448.

[95] T. S. Y. Tang, R. E. Ellis, and G. Fichtinger, *Fiducial registration from a single X-ray image: A new technique for fluoroscopic guidance and radiotherapy*, in Proc. MICCAI 2000, S. L. Delp, A. M. DiGioia, and B. Jaramaz, eds., 2000, pp. 502–511.

[96] T. Theussl, R. F. Tobler, and E. Gröller, *The multi-dimensional Hartley transform as a basis for volume rendering*, in Proc. 8th Int. Conf. Comp. Graph. Visual. Interact. Dig. Med.'2000, vol. 1, 2000, pp. 132–139.

[97] P. Thévenaz, T. Blu, and M. Unser, *Interpolation revisited*, IEEE Trans. Med. Imag., 19 (2000), pp. 739–758.

[98] P. Thévenaz, U. E. Ruttimann, and M. Unser, *A pyramid approach to subpixel registration based on intensity*, IEEE Trans. Image Processing, 7 (1998), pp. 27–41.

[99] P. Thévenaz and M. Unser, *Precision isosurface rendering of 3-D image data*, IEEE Trans. Image Processing, 12 (2003), pp. 764–775.

[100] D. Tomaževič, B. Likar, and F. Pernuš, *Rigid 2D/3D registration of intraoperative digital X-ray images and preoperative CT and MR images*, in Proc. SPIE. Medical Imaging 2002: Image Processing, M. Sonka and J. M. Fitzpatrick, eds., vol. 4684, 2002, pp. 507–517.

[101] V. M. Unger, *Electron cryomicroscopy methods*, Current Opinion in Structural Biology, 11 (2001), pp. 548–554.

[102] M. Unser, *Splines: A perfect fit for signal and image processing*, IEEE Signal Processing Mag., 16 (1999), pp. 22–38.

[103] ——, *Sampling—50 years after Shannon*, Proc. IEEE, 88 (2000), pp. 569–587.

[104] M. Unser, A. Aldroubi, and M. Eden, *B-spline signal processing: Part I—Theory*, IEEE Trans. Signal Processing, 41 (1993), pp. 821–833.

[105] ——, *B-spline signal processing: Part II—Efficient design and applications*, IEEE Trans. Signal Processing, 41 (1993), pp. 834–848.

[106] ——, *The L_2 polynomial spline pyramid*, IEEE Trans. Pattern Anal. Machine Intell., 15 (1993), pp. 364–379.

[107] M. Unser, A. Aldroubi, and S. J. Schiff, *Fast implementation of the continuous wavelet transform with integer scales*, IEEE Trans. Signal Processing, 42 (1994), pp. 3519–3523.

[108] M. Unser, B. L. Trus, and A. C. Steven, *A new resolution criterion based on spectral signal-to-noise ratios*, Ultramicroscopy, 23 (1987), pp. 39–52.

[109] M. Unser, M. J. Vrhel, J. F. Conway, M. Gross, P. Thévenaz, A. C. Steven, and B. L. Trus, *Resolution assessment of 3D reconstructions by spectral signal-to-noise ratio*, in 11th European Congress on Microscopy, Dublin, Ireland, 1996, pp. 260–261.

[110] P. A. van den Elsen, E.-J. D. Pol, and M. A. Viergever, *Medical image matching-a review with classification*, IEEE Engineering in Medicine and Biology Magazine, 12 (1993), pp. 26–39.

[111] M. van Heel, B. Gowen, R. Matadeen, E. V. Orlova, R. Finn, T. Pape, D. Cohen, H. Stark, R. Schmidt, M. Schatz, and A. Patwardhan, *Single-particle electron cryo-microscopy: Towards atomic resolution*, Quart. Rev. Biophys., 33 (2000), pp. 307–369.

[112] M. van Heel and J. Hollenberg, *The stretching of distorted images of two-dimensional crystals*, in Electron Microscopy at Molecular Dimensions, W. Baumeister, ed., Springer, Berlin, 1980, pp. 256–260.

[113] J. A. Velázquez-Muriel, C. O. S. Sorzano, J. J. Fernández, and J. M. Carazo, *A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjustment*, Ultramicroscopy, 96 (2003), pp. 17–35.

[114] R. H. Vogel and S. W. Provencher, *Three-dimensional reconstruction from electron micrographs of disordered specimens. II. Implementation and results*, Ultramicroscopy, 25 (1988), pp. 223–240.

[115] T. J. Wang and T. W. Sze, *The image moment method for the limited range CT image reconstruction and pattern recognition*, Pattern Recognition, 34 (2001), pp. 2145–2154.

[116] J. Weese, R. Göcke, G. P. Penney, P. Desmedt, T. M. Buzug, and H. Schumann, *Fast voxel-based 2D/3D registration algorithm using a volume rendering method based on the shear-warp factorization*, in Proc. SPIE Image Processing 1999, vol. 3661, 1999, pp. 802–810.

[117] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. G. Hill, and D. J. Hawkes, *Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery*, IEEE Trans. Inform. Technol. Biomedicine, 1 (1997), pp. 284–293.

[118] S. Zuffi, A. Leardini, F. Catani, S. Fantozzi, and A. Cappello, *A model-based method for the reconstruction of total knee replacement kinematics*, IEEE Trans. Med. Imag., 18 (1999), pp. 981–991.

Curriculum vitæ

December 4, 2003

Slavica JONIĆ
LIB IOA STI
Swiss Federal Institute of Technology
CH-1015 Lausanne
Switzerland
+41-(21)-693 5136 (work)
Slavica.Jonic@epfl.ch, sjonic@email.com
http://bigwww.epfl.ch/jonic

Born on 7^{th} December 1970 in Negotin, Serbia and Montenegro. Female.

Education

1999 – Present Ph.D. candidate in Biomedical Image Processing at *Swiss Federal Institute of Technology Lausanne (EPFL)*, Switzerland. Thesis title: *"Spline Projection-Based Volume-to-Image Registration"*. Advisor: Prof. M. Unser.

1996 – 1999 M.Sc. in Electrical Engineering at *Faculty of Electrical Engineering, University of Belgrade*, Serbia and Montenegro. Thesis title:*"New Methodology for Obtaining a Rule Base for Bipedal Locomotion Control"*. Supervisor: Prof. D. Popović.

1989 – 1996 B.Sc. in Electrical Engineering at *Faculty of Electrical Engineering, University of Belgrade*, Serbia and Montenegro. Diploma

project in Biomedical Engineering. Supervisor: Prof. D. Popović.

Teaching Experience

1999 – Present Teaching assistant at *Swiss Federal Institute of Technology Lausanne (EPFL)*. Assistant to: *Image Processing and Image Processing Laboratory courses.*
Supervision of semester student projects.

1996 – 1999 Teaching assistant at *Faculty of Electrical Engineering, University of Belgrade*, Serbia and Montenegro. Assistant to: *Electrical Measurements, Biomedical Engineering, and Biomedical Instrumentation.*

Research interests

Signal and image processing, volume-to-image registration, algorithm theory, control theory, machine learning techniques, modelling, numerical methods.

Awards

• Honorable Mention Poster Award at *SPIE Medical Imaging 2003 Symposium: Image Processing*, February 2003, San Diego, California.

Language skills

Mother tongue Serbian, fluent English and French, intermediate German.

Computer skills

Languages: C/C++, Java, Pascal, Fortran.
Packages: Matlab, Simulink, LabView, IDL.
Software: MS Office, Corel Office, LaTeX, Illustrator, Photoshop, ImageJ.
Systems: Mac OS and Windows.

Publications

Journal papers

- S. Jonić, T. Janković, V. Gajić, and D. Popović, *Three machine learning techniques for automatic determination of rules to control locomotion*, IEEE Trans. Biomed. Eng., 46 (1999), pp. 300–310.
- S. Jonić and D. Popović, *Identification of rules for a rule-based control of walking*, Facta Universitatis, Series: Electronics and Energetics, 11 (1998), pp. 1 –24.
- S. Jonić, C. O. S. Sorzano, P. Thévenaz, C. El-Bez, S. De Carlo, and M. Unser, *Spline-based image-to-volume registration for three-dimensional electron microscopy*, Ultramicroscopy. (submitted).
- S. Jonić, P. Thévenaz, G. Zheng, L. -P. Nolte, and M. Unser, *Spline-based registration of a 3D CT to a set of C-arm images*, IEEE Trans. Med. Imag. (submitted).
- M. Jovović, S. Jonić, and D. Popović, *Automatic synthesis of synergies for control of reaching – hierarchical clustering*, Medical Eng. & Phys., 21 (1999), pp. 329–341.
- D. Popović, R. B. Stein, M. N. Oguztoreli, M. Lebiedowska, and S. Jonić, *Optimal control of walking with functional electrical stimulation: a computer simulation study*, IEEE Trans. Rehabil. Eng., 7 (1999), pp. 69–79.
- D. Popović, J. J. Struijk, T. Sinkjaer, and S. Jonić, *Control of near-ballistic walking in paraplegics*, J. Automatic Control, 10 (2001).
- C. O. S. Sorzano, S. Jonić, C. El-Bez, J. M. Carazo, S. De Carlo, P. Thévenaz, and M. Unser, *A multiresolution approach to pose assignment in 3-D electron microscopy of single particles*, J. Struct. Biol. (submitted).

Conference papers

- S. Jonić and D. Popović, *Machine learning for prediction of muscle activations for a rule-based controller*, in Proc. IEEE Ann. Conf. on EMBS, 1997, pp. 1781–1784.

- ——, *Rule-based controller for locomotion - use of radial basis function ANN*, in Proc. Neural Network Applications in Electrical Engineering-NEUREL, 1997, pp. 49–52.

- S. Jonić, D. Popović, and S. Obradović, *Radial basis function neural networks for determining synergies between joint angles in locomotion*, in Proc. Ann. Conf. on Information Technologies, 1998, pp. 124–127.

- S. Jonić, P. Thévenaz, and M. Unser, *Multiresolution spline-based 3D/2D registration of CT volume and C-arm images for computer-assisted surgery*, in Proc. SPIE. Medical Imaging 2001: Image Processing, M. Sonka and K. M. Hanson, eds., vol. 4322, 2001, pp. 1101–1109.

- ——, *Multiresolution-based registration of a volume to a set of its projections*, in Proc. SPIE. Medical Imaging 2003: Image Processing, M. Sonka and J. M. Fitzpatrick, eds., vol. 5032, 2003, pp. 1049–1052.

- D. Popović and S. Jonić, *Identification of one-dimensional biomedical signals: Supervised machine learning techniques*, in Proc. Neural Network Applications in Electrical Engineering-NEUREL, 1997, pp. 182–187. (invited paper).

- ——, *Determining synergy between joint angles during locomotion by radial basis function neural networks*, in Proc. IEEE Ann. Conf. on EMBS, vol. 5, 1998, pp. 2301–2304.

- ——, *Control of bipedal locomotion assisted with functional electrical stimulation*, in Proc. American Control Conference, 1999, pp. 1238–1242.