11-13-2009

# Rigid and Non-rigid Point-based Medical Image Registration

Nestor Andres Parra
*Florida International University*, nparr003@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

RIGID AND NON-RIGID POINT-BASED MEDICAL IMAGE REGISTRATION

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER SCIENCE

by

Nestor Andres Parra

2009

To: Dean Amir Mirmiran
      College of Engineering and Computing

This dissertation, written by Nestor Andres Parra, and entitled Rigid and Non-rigid Point-based Medical Image Registration, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Anthony McGoron

_____
Geoffrey Smith

_____
Tao Li

_____
Giri Narasimhan, Major Professor

Date of Defense: November 13, 2009

The dissertation of Nestor Andres Parra is approved.

_____
Dean Amir Mirmiran
College of Engineering and Computing

_____
Dean George Walker
University Graduate School

Florida International University, 2009

DEDICATION

To Carlos A. and Maria L., in gratitude for the gift of life.

To Mark J., Bryan P., Shira K., and those who carry the message.

ABSTRACT OF THE DISSERTATION

RIGID AND NON-RIGID POINT-BASED MEDICAL IMAGE REGISTRATION

by

Nestor Andres Parra

Florida International University, 2009

Miami, Florida

Professor Giri Narasimhan, Major Professor

The primary goal of this dissertation is to develop point-based rigid and non-rigid image registration methods that have better accuracy than existing methods. We first present point-based PoIRe, which provides the framework for point-based global rigid registrations. It allows a choice of different search strategies including (a) branch-and-bound, (b) probabilistic hill-climbing, and (c) a novel hybrid method that takes advantage of the best characteristics of the other two methods. We use a robust similarity measure that is insensitive to noise, which is often introduced during feature extraction. We show the robustness of PoIRe using it to register images obtained with an electronic portal imaging device (EPID), which have large amounts of scatter and low contrast. To evaluate PoIRe we used (a) simulated images and (b) images with fiducial markers; PoIRe was extensively tested with 2D EPID images and images generated by 3D Computer Tomography (CT) and Magnetic Resonance (MR) images. PoIRe was also evaluated using benchmark data sets from the blind retrospective evaluation project (RIRE). We show that PoIRe is better than existing methods such as Iterative Closest Point (ICP) and methods based on mutual information.

We also present a novel point-based local non-rigid shape registration algorithm. We extend the robust similarity measure used in PoIRe to non-rigid registrations adapting it to a free form deformation (FFD) model and making it robust to local

minima, which is a drawback common to existing non-rigid point-based methods. For non-rigid registrations we show that it performs better than existing methods and that is less sensitive to starting conditions. We test our non-rigid registration method using available benchmark data sets for shape registration.

Finally, we also explore the extraction of features invariant to changes in perspective and illumination, and explore how they can help improve the accuracy of multi-modal registration. For multimodal registration of EPID-DRR images we present a method based on a local descriptor defined by a vector of complex responses to a circular Gabor filter.

TABLE OF CONTENTS

LIST OF FIGURES

# CHAPTER 1

## INTRODUCTION

*Image Registration* is the process of finding the spatial alignment between two images. Fitzpatrick defines it as "the determination of a geometrical transformation that aligns points in one view of an object with corresponding points in another view" [FWM98]. The term *view* is used to include images as well as real 3D objects as in patient registration for image guided surgery. In a medical imaging context, it was defined by Maurer as "the determination of a one-to-one mapping between the coordinates in one space and those in another such that points in the two spaces that correspond to the same anatomical point are mapped to each other" [MF93]. Zitova emphasizes that the views to be registered could have been taken at different times, from different view points, using different equipment (multi-modal), or could even be on different subjects (inter-patient) [ZF03].

In computer vision, image registration has a wide range of applications such as image matching for stereo vision, motion analysis, three-dimensional shape estimation and object identification (pattern recognition). In medical imaging, image registration is required for applications such as multi-modality fusion, image segmentation, deformable atlas registration, functional brain mapping, image guided surgery, medical diagnosis, treatment planning, characterization of normal versus abnormal anatomical shape and variation (progress of disease or result of treatment) and to guide minimally invasive medical procedures and treatments.

The images to be registered can be acquired using diverse types of equipment such as satellite imaging, range imaging or infrared thermography. Medical images differ from other types of imaging mainly in three aspects: a) Medical images emphasize the analysis of internal structures in the subject. b) Medical images of different subjects are, in general, expected to differ substantially, and c) Medical images of

(a) EPID image        (b) CT slice        (c) MR slice
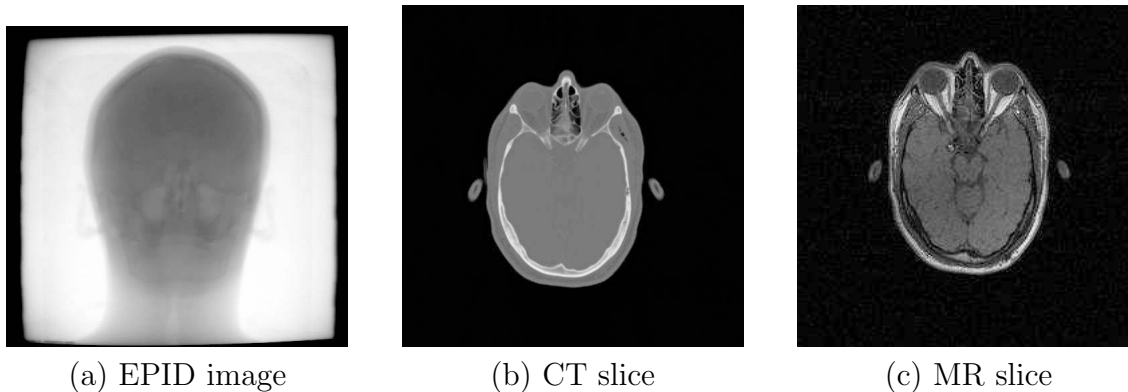
Figure 1.1: A few medical imaging modalities

the same subject differ substantially during growth or in the presence of abnormalities. Modalities in medical imaging include Computed tomography (CT), Magnetic resonance Imaging (MRI), Electronic portal imaging (EPID), and ultrasound. All of these provide an anatomical/physiological description of the patient (useful for trauma assessment, radiotherapy planning, etc.). Alternatively, Positron Emission Tomography (PET), functional MRI (fMRI) and spectroscopy provide functional information such as blood flow, oxygen use, and glucose metabolism. These methods are able to measure functional information (tumor growth by measuring levels of glucose absorption, neurological activity by measuring the levels of certain neurotransmitter, etc.). Figure 1.1 shows three examples of different medical images. 3D modalities are usually represented as a stack of 2D images. We are particularly interested in image registration in radiotherapy where one important application is to verify patient position with respect to the coordinate system of the medical linear accelerator (or LINAC) used prior to radiation beam delivery as part of treatment verification. Treatment outcome is intrinsically linked to precise beam delivery and patient positioning, which are needed to maximize radiation dose to the tumor and to minimize its effects on neighboring critical and normal anatomical structures.

Image registration is used to achieve accurate patient positioning.

Before we present the different aspects of image registration it is important to point out the differences between point sets and images. In the context of imaging, a pixel (picture element) is a sample of an original continuous view. The sampling frequency determines the pixel size. The smaller the pixel-size, the more accurate is the representation. Pixels represent 2D image data in in a spatially mapped array of bits (bitmap). The term *voxel* is used for a 3D pixel (volumetric pixel). Each pixel has intensity information associated with it. The number of bits allowed to represent these intensities indicate the level of granularity in the image. In gray-scale images a single value represents the intensity. In color images using the RGB representation, there are three values (channels) for red, green and blue intensities for each pixel. In this dissertation we only use gray-scale images. In general, pixels are arranged in a 2-dimensional grid and each pixel is not described by its coordinates (which could be inferred from the pixel position in the array). In contrast, points are represented by their coordinates. They can be considered an abstraction of a pixel and do not have a size associated with them. In the image registration context, this is a subtle point. When we apply a geometric transformation to the pixels in an image, this operation could potentially map pixels to locations outside the integer domain and assign pixel locations to points in the real plane. Thus transformed pixels have to be rounded off to the nearest (new) pixel location. Since this introduces signal noise, often the contribution of neighboring pixels is measured and the new pixel intensity value is computed by an averaging process. Consequently, after the geometric transformation the intensities of pixels change, affecting the performance of voxel-based methods (methods that use pixel intensities to register images). When a geometric transformation is applied to a point, their coordinates are modified but

(a) original pixel image

(b) rotated pixel image

(c) extracted point set
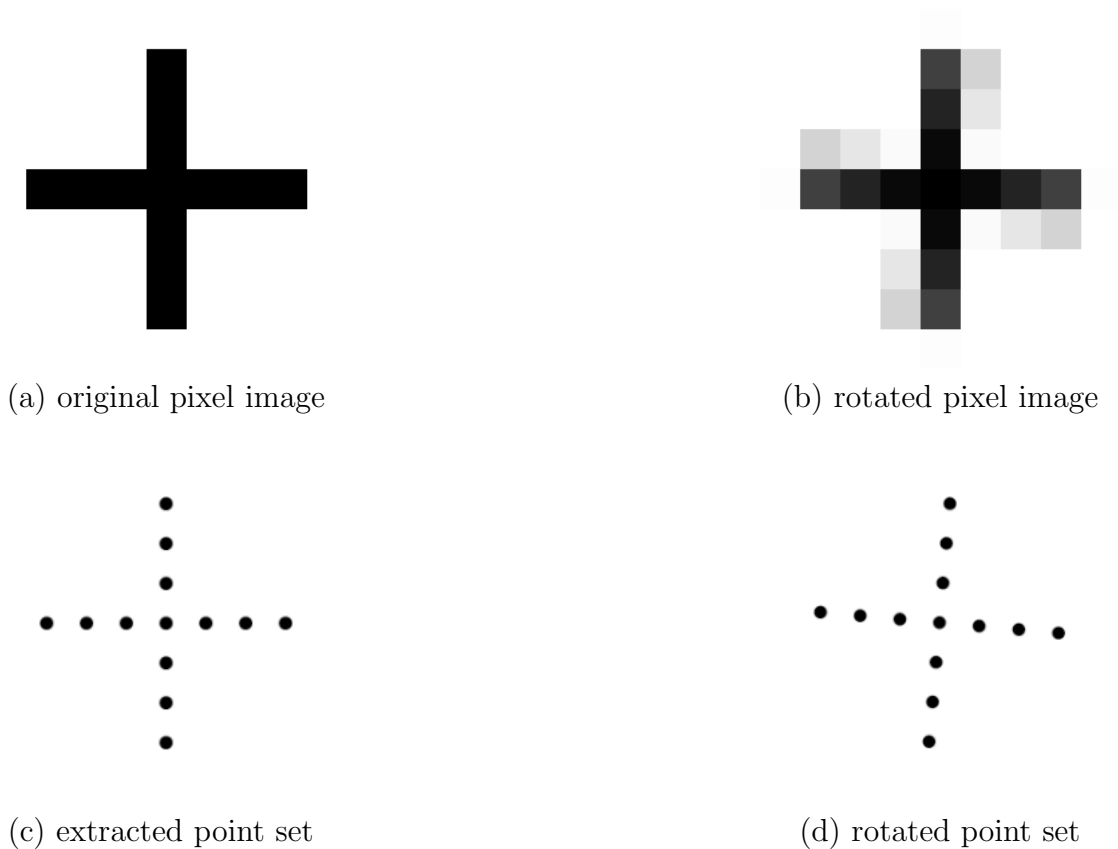
(d) rotated point set

Figure 1.2: Rotation of pixels and points

theirs associated values (intensity, gradient, etc) are usually left unchanged.

Figure 1.2 shows the difference between applying a rotation to an image and to a point set. The image size is $9 \times 9$, meaning that the grid has 9 rows and 9 columns. After the small rotation the pixel intensities of the pixels close to the original cross pattern get interpolated and the resulting image is not exactly the rotated image. On the other hand, the point set representation is able to maintain the integrity of the image even under geometric transformations.

We define a *shape* as the binary representation of an object. In other words, the intensity values are binary (0 or 1). The process of feature extraction is trivial with shapes and we consider them "simple" images. When testing our algorithm for the challenging problem of non-rigid registration, most of our testing is confined to the

registration of shapes instead of the more general images. However, we will show that registration of shapes can be extended to image registration by extracting the features of interest in the image and applying the shape registration algorithm to the extracted point features (referred to as *edges*.

Image registration has been the subject of extensive research. From an algorithmic perspective, Brown [Bro92] has classified image registration methods according to the following criteria:

1. **Feature space:** It is defined by the representation of the input and corresponds to the information extracted from the images to guide the registration. Maintz [MV98] classified image registration methods into two broad classes: voxel-based [HF96, KFL$^+$01, MCV$^+$97, MVS99, PST$^+$00, SHF99, TU00] and feature-based [BPC92, GvH93, GTvHV95, LLB98]. In voxel-based methods, the image intensity values are directly used to compute and evaluate the registration. On the other hand, feature-based methods rely on a feature extraction step and features are used to perform the matching. Typical features of interest include edges, contours, and boundaries that are represented by points, curves, surfaces, volumes, etc. In medical images, these features may correspond to the bone anatomy of the patient. Audette et al. make a distinction between different feature-based representations: *Point features* are sparse, salient, well-localized loci of important *local* significance. In contrast, dense point-based schemes, referred to as *point clouds* or free-form surface registration, are sets of points coordinates without additional information [AFP00]. Methods based on invariant local descriptors try to determine correspondences between points based on local information. The *scale-invariant feature transform (SIFT)*, introduced by Lowe, is invariant to similarity transformations thanks to a relative point localization in scale and orientation [Low99]. Freeman introduced

the *steerable filters* that are computed by steering the derivatives in a particular direction given the components of a set of local derivatives (*local jet*) [FA91]. Steering derivative in the direction of the gradient makes them invariant to rotation. Moment invariants were introduced by Van Gool to describe the multispectral nature of data by using central moments of low order and low degree [GMU96]. For images that are shapes, they are often represented by transforms such as the distance transform. The transformed images are then used for the registration.

2. **Transformation space:** It corresponds to the representation of the transformation used to align the images and defines the search space that has to be explored to solve the problem. In the global case it corresponds to the space of valid rigid, affine or perspective transformations. Local non-rigid transformation models include optical flow, Thin Plate Splines (TPS) and Free Form Deformations (FFD). Transformations can be either global or local. Global registration aims to find the parametric transformation (usually rigid, affine or projective) that maps one image as closely as possible to another one maximizing some similarity function. Local registration is needed to account for non-rigid deformations that affect only a part of the image.

3. **Similarity measure**: It is the criterion used to evaluate the similarity between two images and therefore is used to evaluate a given transformation. It is tightly related to the feature space. Examples of similarity measures associated with intensity-based image registration methods include *mutual information* [VW95], *cross-correlation* [Zhu02], and *local frequency representation* [LVM02]. For landmark registrations, a commonly used similarity measure is the sum of squared distances. Other example of feature-based similarity measures are the *chamfer distance* [Bor88], the *Hausdorff distance*, and the *Frechet*

*distance* (which operates on parametrized curves). The Hausdorff distance was proposed in 1993 by Huttenlocher et al. and does not require the computation of pair matching. Huttenlocher et al. introduced the partial Hausdorff distance which additionally provides robustness to outliers [HR93]. In an effort to provide extended robustness toward outliers, Takaks presents a weighted Hausdorff method that uses a neighborhood radius and a penalizing factor to consider only local points. Previous weighted versions of Hausdorff have been proposed. One includes the work by Lin [LLS03] in which the weights favor points with large first eigenvector for a given training set; this corresponds to areas where one should pay more attention for classification, like the eyes and the mouth in a facial image. Although effective for classification, this weight function is not useful for registration, where it would be nice to favor corresponding areas of the face (like ears, or nose) to have small distance between them.

4. **Search strategy:** It corresponds to the strategy used to explore the search space and maximize a similarity measure (or minimize a distance function). For voxel-based methods, it usually corresponds to a variational approach where a gradient descent-like method is used to minimize an energy functional. Gradient descent methods are iterative methods, generally sensitive to starting conditions and can get trapped in local minima. Point-based methods usually require a determination of the point correspondence between the point sets followed by a cost minimization between established correspondences. A potential drawback of feature-based registration algorithms is that the accuracy is limited by the quality of the feature extraction. Hierarchical techniques are also used and aim to improve and speed up the registration by performing the search through progressively finer resolutions. The HAMMER algorithm
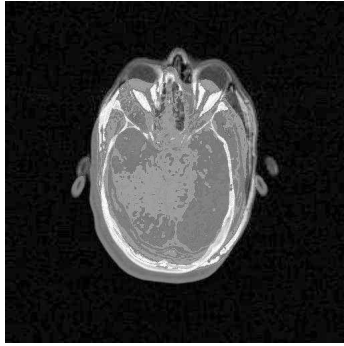
(a) CT and MR before registration
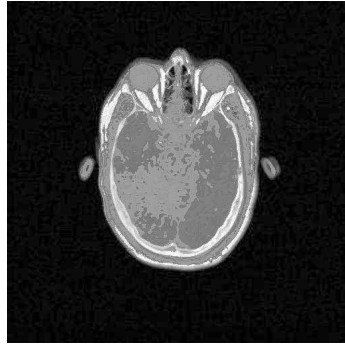
(b) CT and MR after registration

Figure 1.3: CT and MR slices overlapped before and after registration

for elastic registration of brain images presented by Shen minimizes an energy function of a hierarchically growing set of voxels, and is thus both voxel- and feature-based [SD02]. In general, during the process of extraction of interesting features, undesirable or noise features also get enhanced, introducing considerable amounts of "noise" into the point set. For point-based methods the strategy most commonly used is the *iterative closest point ICP* that performs a two-stage minimization of the distance between points.

Given two image representations $A$ and $B$, a transformations space $T$, and a similarity measure $sim$; the image registration problem can be presented as computing the following:

$$t_{OPTIM} = \arg\max_{t \in T} sim(A, t(B)).$$

Figure 1.3 shows the overlap of a CT and MR image before and after registration. In this example we are using a pixel representation for the images and the space of transformations consists of translations only.

This dissertation work focuses on point-based registrations. As mentioned before, points can be represented as point sets (point clouds or graph representation)

or as implicit distance functions. Also, additional information added to each point (local descriptors) may assist the algorithm in improving the registration.

Using edges for image processing dates back to when Biederman showed that humans can identify faces based on the outline as accurately as with photographs [BJ88]. His research suggests that information such as color and texture are not used as indices to retrieve information from memory. Additionally, Glynn shows that simple cortical cells operate on edges instead of full intensity values [Glynn]. These findings motivated several researchers to present methods for object and face recognition using the outline of images [Tak98, SKP99, GS99, YLC04, YLC07]. These outlines are commonly represented as a set of points or pixel locations extracted using different edge detection algorithms, like the Laplacian of Gaussian (LoG) [HM86], Canny [Can86], or odd Gabor filter-based edge detection [ZTL04]. Edges have also been represented by curves and manifolds in 2D and 3D but such approaches have the problem of requiring more complex differential solutions. For the case of edge representation several similarity measures have been proposed to compare two edge sets. These include the Hausdorff distance and several of its variants.

Several algorithms have been proposed for point set registration. Atallah was one of the first to study this problem in a theoretical setting [Ata85]. His $O(n \log^2 n)$ algorithm checks the existence of a one-to-one matching between two point sets of the same size using the property that in a rigid transform-based matching, all the line segments joining matched pairs of points are non-intersecting. The application of this algorithm is limited to cases where both data sets have the same size, which is a very severe restriction, considering the noisy nature of data in practical applications. Other algorithms have been proposed for exact registration. Goodrich used several points as anchor points to find ideal pairs of points to define a transformation

[GMO99]. Their running times were compared to a branch-and-bound method but the experiments used very small data sets. Bishnu proposed an algorithm for partial registration that placed several constraints on the data sets making it also unsuitable for image registration [BDNB06]. Their work shows an interesting partial matching algorithm for point sets of different cardinality where one is a subset of the other. This handles noise points in the reference image but assumes no noise in the target, again an unrealistic assumption. Other restrictions of their method include having the points sparse enough so that balls around the points with radius equal to the size of the desired maximum error do not intersect. Other approximate algorithms for matching of point sets involve computing the intersection of high degree curves that makes them unstable, difficult to implement and have high running times.

Accurate image registration can be performed effectively if some reference (fiducial) markers are available in the images. Where natural markers are not available, artificial ones have often been introduced. Since skin is motile, markers must be placed at rigid anatomical sites through a surgical procedure, making this a potentially invasive procedure, especially if high precision is sought [Mur02]. This makes markers undesirable. Our algorithms specifically avoid assumptions about the availability of markers and are designed to perform well in their absence.

Corresponding landmark-based registration has been used as a general solution for the registration of two sets of points, and widely used for marker-based registration. This approach was first presented by Schonemann [Sch66], and more recently, Umeyama [Ume91] and Kanatani [Kan94] introduced several improvements. Landmark-based methods for point registration involve finding the least-squares estimate for the norm of the difference between a set of points and a transformed one. It is worth mentioning that this method requires that both sets of points have the same size, it is not robust to missing or noisy points and the correspondence between

the point sets has to be provided. Establishing this correspondence requires an *a priori* approximate solution of the registration.

Even though extensive research on point-based registration has been done, there are no clinical studies of point-based registrations other than the ones using marker-based methods. There is no clinical study that evaluates the accuracy and running time of a point-based method using extracted edges. In radiotherapy, there is a need to be able to register portal images accurately with their low visual quality, high level of scatter, in a setting that is robust to out-of-plane rotations. We show in our work that Mutual Information is unstable in the presence of out of plane rotations. This work is mainly oriented to (1) present solid evidence that point-based registration produces accurate registrations even for noisy images, (2) incorporate additional information for the extracted edges such as gradient information, (3) develop a method that weights points based on their relevance and interest, and finally, (4) extend the point-based method to non-rigid registrations.

The **primary aim** of this dissertation *is to explore the field of point-based medical image registrations and present novel methods for rigid and non-rigid registrations, and to study the use of local descriptors for multi-modality registration*s. We aim to develop a point-based method that is relatively insensitive to noise introduced during the feature extraction step, accurate to out-of-plane rotations and integrates additional information for smart registration. The central hypothesis is that robust and accurate image registration of feature-based point sets can be achieved when using the appropriate similarity measures and search strategies. We present the 5 goals of this dissertation below.

<u>**Goal 1:**</u> In medical imaging, point-based registration methods have been limited to landmark-based registration. Although point-based methods can be potentially faster than intensity-based methods, the registration of extracted edge

points has not proved to be more accurate than voxel-based registration. *The first goal of this dissertation is to explore point-based image registration methods and make them more accurate than other existing point-based methods without the use of landmarks.*

**Goal 2:** Current point-based methods require the assignment of a correspondence between points in the two images. This is a drawback for automated methods because computing the correspondence is very sensitive to noise. *Our second goal is to explore new registration schemes that do not depend on computing a point correspondence.*

**Goal 3:** Local descriptors have been the topic of extensive research. These descriptors add information to the point coordinates that can guide the registration. *The third goal is to explore ways to use additional information in improving the registration process.*

**Goal 4:** Multi-modal registration is a challenging topic. It is well known that point-based registration of edges extracted from images of different modalities is subject to error because of the relatively poor spatial correspondence between features in the images. *The fourth goal of our work is to achieve accurate multi-modal registration.*

**Goal 5:** Non-rigid registration is currently the leading area of research in image registration. *The last goal of this dissertation is to explore the suitability of point-based registration to solve the problem of non-rigid registration.*

The remainder of this dissertation is organized as follows.

Chapter 2 presents a review of existing methods used in image registration following the four criteria used in this chapter: feature-space, transformation space, similarity measure and search strategy.

Chapter 3 introduces a point-based registration method for rigid registration. The search strategies will include branch-and-bound search, stochastic hill climbing and a novel hybrid method. We compare the performance of our method to that of using popular methods including mutual information and ICP.

Chapter 4 presents several local descriptors such as gradient magnitude and responses to a complex Gabor filter. The latter has been known to extract texture information. These descriptors are then used to improve the registration of multi-modal imaging.

Chapter 5 focuses on non-rigid registration and deformable models. We show how a modified version of the similarity measure used in Chapter 3 allows the computation of accurate non-rigid registration using a free form deformation model.

We finish with Chapter 6 where we present the conclusions of our work and suggest future directions.

CHAPTER 2

**BACKGROUND**

This chapter describes the necessary background for the work presented in this dissertation. We start by describing the four imaging modalities used in this dissertation. We follow the same four criteria for describing registration algorithms as presented in the previous chapter. After that, we describe similarity measures for voxel-based and point-based registrations. We then describe rigid, affine, and non-rigid transformation spaces. We proceed to describe several search strategies and conclude the chapter with techniques to evaluate the accuracy of point-based registration methods.

## 2.1 Imaging modalities

Although our algorithms are general enough to manipulate any type of imaging, our experiments emphasize four imaging modalities described below.

### 2.1.1 Computed tomography (CT)

Computed tomography (CT) volumes are used during diagnosis and also in patient treatment planning calculations. These kilovoltage X-ray 3D images display good contrast and high resolution. CT images are needed for both anatomical target identification and for modeling radiation beam interactions. Digitally reconstructed radiographs (DRRs) are 2D projections of CT images for a specified source-patient-detector geometry. The CT (and DRR) provide the "reference" images regarding the intended patient position in the sense that the beam planning assumes a specific alignment of the patient with the LINAC coordinate system. The process of aligning the patient in the treatment couch with the CT volume is called patient verification.

Typically, verification of patient position is carried out by 2D portal imaging immediately prior to radiotherapy treatment using megavoltage X-rays. Patient position can be verified either by comparing successive portal images, or by comparing the portal images with previously captured DRRs.

## 2.1.2 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging (MRI, or MR for short) has several advantages over X-ray CT imaging. It produces images with better contrast for soft tissue and has the ability to directly obtain images in any plane allowing an easier segmentation of tumors while providing a clearer delineation of certain tissue boundaries. It has also several limitations: it is more susceptible than CT to spatial distortions and intensity artifacts, lacks signal from bone, and produces pixel values that have no relationship to electron densities. MRI is now established as an important complement to CT imaging for treatment planning.

## 2.1.3 Positron Emission Tomography (PET)

Positron emission tomography, also called PET imaging or a PET scan, is a type of nuclear medicine imaging. Nuclear medicine is a branch of medical imaging that is noninvasive and uses small amounts of radioactive material to diagnose or treat a variety of diseases, including many types of cancers, heart disease, and certain other abnormalities within the body. These imaging scans use radioactive materials called radiotracers. Depending on the type of nuclear medicine exam, the radiotracer is either injected into a vein, swallowed or inhaled as a gas and eventually accumulates in the organ or area of the body being examined, where it gives off energy in the form of gamma rays. This energy is detected by a device called a gamma camera,

a (positron emission tomography) PET scanner and/or probe. These devices work together with a computer to measure the amount of radiotracer absorbed by the body and to produce special pictures offering details on both the structure and function of organs and tissues.

### 2.1.4   Electronic portal imaging device (EPID)

Portal imaging is acquired using megavoltage (same energy used for treatment in radiotherapy) X-rays with an electronic portal imaging device (EPID), which has the advantages of online access and reduced exposure dependence, as well as the usual advantages of digital imaging over film. The main advantage of built-in EPID systems is that they provide an accurate spatial localization of the patient position based on bony anatomy. The main reason for relatively poor spatial and contrast resolution of EPID imaging compared to radiographic (i.e., kilovoltage) imaging is the dominance of Comptom scattering and low detective quantum efficiency of EPIDs.

## 2.2   Feature Space

Images are commonly represented as an array of voxels. However, images and shapes are often best described in terms of their features. Feature extraction methods help to extract important and relevant features from images. Of particular interest are *local descriptors*, which extract invariant information from the image, are discriminant in the selection of points, and are used in several registration schemes.

### 2.2.1 Comparative discussion

Experimentally, Canny's algorithm ([Can86]) produced the best results for EPID images. The Laplacian of Gaussian (LoG) algorithm [HM86] was chosen for DRR images. For the CT and MR images we used the Deriche algorithm ([Der93]) which uses the derivatives of a Gaussian filter for edge detection.

### 2.2.2 Voxel-based representation

As we mentioned in Chapter 1, digital images have discrete representations. Images sample a subject at a finite number of pixels. For voxel-based representation, the domain $\Omega$ is defined as $\Omega =: \tilde{\Omega} \cap \Gamma_\zeta$, where $\tilde{\Omega}$ is a *bounded* continuous set defining the volume of the subject and $\Gamma$ is the infinite discrete grid with a particular pixel size $\zeta$. Given two images $A$ and $B$, the sampling is in general different from each image and defines two separate domains $\Omega_A$ and $\Omega_B$.

### 2.2.3 Point set representation

Feature extraction methods transform binary images into dense sets of points, often referred to as point clouds. A point set can be represented simply as a list of nodes or also as a *Distance transform* (DT), which has been shown to be an efficient representation of the point set. Distance transforms convert a binary image into a non-binary image where each pixel value denotes the distance to the nearest non-zero pixel within the shape.

A distance transform embeds a given point set $B$ as the zero level set of an

(a) Point set representation



(b) Distance Transform representation
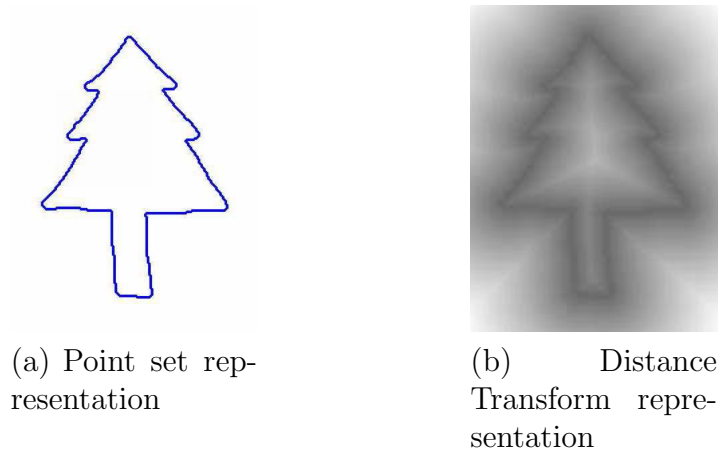
Figure 2.1: Point set representation

unsigned distance function defined in 2D as follows:

$$\Phi_B(X) = \Phi_B(x, y) = \begin{cases} 0, & (x, y) \in B \\ D((x, y), B) > 0, & (x, y) \notin B \end{cases}$$

where $D((x, y), B)$ corresponds to the minimum distance between the point $(x, y)$ and the points in $B$. The algorithm by Felzenszwalb and Huttenlocher implements the minimum Euclidean distance [FH04].

The *chamfer distance*, approximates the Euclidean distance transform by counting the number of horizontal/vertical and diagonal moves needed for a pixel to reach the shape. This can be done by propagating local weighted distances. It can be computed efficiently with two raster scans of the binary image. Since we compute the distance transform only once at the beginning of the algorithm, we can afford the more expensive computation of the Euclidean distance transform.

Figure 2.1 shows two point set representations. For the points shown as a set of points in Figure 2.1(a), the Euclidean distance transform representation is shown in Figure 2.1(b).

**Local descriptors**

Local descriptors are methods that describe local image features that are relatively invariant to changes in illumination, changes in viewpoint (invariant to rigid, affine, or perspective transformations), and are robust to clutter and partial occlusion. At the same time, local features must be distinctive enough so that different objects can be identified. This is particularly important when using local descriptors to guide the point-based image registration because the objects with a correspondence in space need to have highly correlated local features.

The *Scale Invariant Feature Transform (SIFT)* presented by Lowe [Low99] has four steps. First, identification of important locations in scale space is performed by computing the extremes (*keypoints*) of a *Difference-of-Gaussian (DOG)* function. This in done by creating a pyramid of images subsampled at different rates and looking for such keypoints in all images. The second step localizes these keypoints to subpixel accuracy and discards them if unstable. The third step computes a relative orientation for each located point based on the dominant orientation of the area around that point. This relative localization in scale and orientation allows the SIFT filter to be invariant to similarity transformations. The fourth step builds a local image descriptor of a point based upon the image gradients in its local neighborhood.

Belongi et al. introduced a very discriminative local descriptor called *shape context* that computes a histogram describing the edge distribution in a region [BMP01]. The authors use a log-polar histogram to represent edge location relative to each feature point that has 12 bins for rotation and 5 bins for the log of the radius.

A local descriptor based on moment invariants was introduced by Van Gool to describe the multispectral nature of data [GMU96]. The invariants use central mo-

ments defined by $M_{pq}^a = \iint_\Omega x^p y^q \left[I\left(x, y\right)\right]^a dx\, dy$ of order $p + q$ and degree $a$. The authors used up to the second order and second degree moments because moments of higher order and degree are sensitive to small geometric and capture distortions. These descriptors have been found to be more suitable for color images where invariants can be computed for each color channel and between channels [MLS05].

Several authors have used the properties of local derivatives (*local jet*) to create local descriptors. A stable estimation of the derivative can be approximated by convolution with Gaussian derivatives. Freeman presented a local descriptor based on steerable filters computed by steering derivatives in a particular direction given the components of the local jet [FA91]. Steering derivatives in the direction of the gradient makes them invariant to rotation.

## 2.3  Similarity Measures

### 2.3.1  Voxel-based Similarity Measures

When using voxel intensities to guide the registration of two images $A$ and $B$ and we want to evaluate how well a transformation $T$ aligns the two images, we apply the transformation $T$ to the target image to produce a mapped image $B^\tau$. Note that this process involves an interpolation function to fit the values of $B^\tau$ to the overlap domain $\Omega_{A,B}^T$ used in voxel-based similarity measures and defined as $\Omega_{A,B}^T = \left\{x_A \in \Omega_A \,|\, T^{-1}(x_A) \in \Omega_B\right\}.$

**Sum of squared intensity differences (SSD)**

The basic voxel-based similarity measure uses the voxel intensity differences

$$SSD = \frac{1}{N} \sum_{x_A \in \Omega_{A,B}^{\mathbf{T}}} \left|A(X_A) - B^\tau(X_A)\right|^2.$$

This is the optimum similarity measure when the images differ only by Gaussian noise. While this situation happens for intramodality registration if the noise from acquisition equipment is Gaussian, it is very unlikely to hold for inter-modality registration. Also, this similarity measure is not robust since it can be dominated by a small number of pixels having large intensity differences. Another similarity measure uses the sum of absolute differences (SAD) and is defined as follows:

$$SAD = \frac{1}{N} \sum_{x_A \in \Omega_{A,B}^{\mathbf{T}}} |A(X_A) - B^\tau(X_A)| \,.$$

The SAD measure is more robust than the SSD measure since it does not amplify differences as much as the SSD measure. More robust non-linear measures are possible if median is used instead of averaging.

## Cross correlation

Cross correlation assumes that there is a linear relationship between the intensity values of the images. Again, this does not hold for inter-modality registration. It is defined as follows:

$$CC = \frac{\sum_{x_A \in \Omega_{A,B}^{\mathbf{T}}} (A(X_A) - \bar{A}) \cdot (B^\tau(X_A) - \bar{B})}{\sqrt{\sum_{x_A \in \Omega_{A,B}^{\mathbf{T}}} (A(X_A) - \bar{A}) \cdot \sum_{x_A \in \Omega_{A,B}^{\mathbf{T}}} (B^\tau(X_A) - \bar{B})}} \,.$$

It is more robust than the SSD measure but the assumption of linearity limits its application.

## Mutual Information

Mutual Information (MI) [VW95] has become the most popular voxel-based similarity measure for image registration mostly because of its success in inter-modality registration. It uses the concept of entropy from information theory and measures the amount of shared information between two images.

The mutual information $I(A, B)$ between two images $A$ and $B$ is defined as

$$I(A, B) = H(A) + H(B) - H(A, B),$$

where $A$ and $B$ are random variables associated with the pixel intensities. These random variables are defined using the histogram of each image. $H(A)$ and $H(B)$ are their marginal entropies. The entropy of a random variable $S$ is defined as

$$H(S) = -\sum_s p_S(s) \log p_S(s),$$

where $p_S(s)$ is the marginal distribution. The joint entropy $H(A, B)$ is defined as

$$H(A, B) = -\sum_{a,b} p_{AB}(a, b) \log p_{AB}(a, b),$$

where $p_{AB}(a, b)$ is the joint probability distribution.

We used the multi-resolution MI approach by Thevenaz, which uses two traditional methods for optimization ([Pow64] and [Mar63]); the joint histogram is estimated by quasi-random sampling and B-spline is used as the Parzen window to estimate the joint histogram. Their interpolation model is a cubic spline [TU00].

Voxel-based registration has become the method of choice in medical image registration. Mutual information has successfully addressed the problem of multi-modal image registration for some modalities like CT and MR. In radiotherapy applications, mutual information is not robust for registering EPID images. For this reason, the correlation coefficient has been used for multi-modal EPID-DRR registration. It assumes a linear relationship between the two modalities, which does not hold due to the physics of high energy imaging. Voxel-based registration is popular in medical imaging because point-based methods require a feature extraction step that can potentially introduce noise and affect the quality of the registrations. One of the objectives of this dissertation is to explore robust point-based registration methods and evaluate their accuracy.

## 2.3.2  Point-based similarity measures

**Chamfer matching**

In point-based registration methods, the chamfer distance is commonly used. Given two point sets $T$ and $I$, it is defined as $D_G(T, I) = \frac{1}{|T|} \sum d_I(T)$ [Bor88]. The root mean square distance has also been used as the matching measure and defined as $D_{rms} = \frac{1}{3} \sqrt{\frac{1}{n} \sum_{i=1}^{n} v_i^2}$. Algorithms that compute distance transforms based on the exact Euclidean distance between points are available and have a time complexity that is linear in the total number of pixels/voxels in the image.

Although chamfer distance only approximates the Euclidian distance, it has been widely used under the assumption that feature extraction is a process that introduces uncertainty and noise in the point locations and therefore there is no need to measure Euclidean distances exactly during feature-based registration. We used the exact Euclidian distance transform because both the exact and approximate distance transform computations are linear in the size of the image and the performance was not affected by using the exact method.

**Hausdorff distance**

The Hausdorff distance has been used for a long time because it does not require the explicit computation of matching pairs. Huttenlocher [HR93] introduced the partial Hausdorff distance which additionally provides robustness to outliers. In an effort to provide additional robustness in the presence of outliers, Takaks presented a weighted Hausdorff method that uses a modified Hausdorff distance using a neighborhood radius and penalizing factor to consider only local points [Tak98].

In what follows, we formalize the definition of the Hausdorff distance and its variants. The Hausdorff distance does not require that both sets have equal number

of points. Given two point sets $A$ and $B$, the Hausdorff distance from $A$ to $B$ is defined as

$$H(A, B) = \max(H_D(A, B), H_D(B, A)).$$

where $H_D(A, B)$ is the directed (or asymmetric) Hausdorff distance and is defined as

$$H_D(A, B) = \max_{a \in A} \min_{b \in B} dist(a, b)$$

where $dist(a, b)$ is any distance metric between two points. (Note that the Euclidean distance is the most commonly used measure of distance between two points.)

A more "robust" measure is that of the partial Hausdorff distance defined as

$$H_k(A, B) = \overset{k}{\max_{a \in A}} \min_{b \in B} dist(a, b), \tag{2.1}$$

where $\max_{a \in A}^k$ returns the $k^{th}$ smallest element of the set of values obtained for $a \in A$. Note that $\max_{a \in A}^{|A|} = \max_{a \in A}$ This measure is robust because it requires that only $k$ pairs of points be matched, and is insensitive to the magnitude of the errors for the unmatched points. The value $k$ is usually given as a quantile $q$, with $0 < q \le 1$ and $k = q |A|$. Henceforth, we will refer to $q$ as the *robustness quantile*. The modified Hausdorff distance is robust to noise introduced during the feature extraction. We will refer to the partial Hausdorff distance as the *unweighted Hausdorff distance*.

We improve on the partial Hausdorff distance by introducing a *weighted* version of the above distance measure, which weights the points differently (and correspondingly the errors in their matching).

Debuisson and Jain [DJ94] defined a modified Hausdorff distance as follows:

$$H_{DJ}(A, B) = \frac{1}{N_a} \sum_{a \in A} \min_{b \in B} \|a - b\| \, .$$

Takaks introduces another variant (M2HD) defined as follows:

$$H_T(A, B) = \frac{1}{N_a} \sum_{a \in A} d(a, B),$$

where

$$d(a, B) = \max(I \min_{b \in N_B^a} \|a - b\|, (1 - I)P)$$

and $I$ is a binary variable with value 1 if the *neighborhood function* set $N_B^a$ is not empty and 0 otherwise [Tak98].

Wang et al. defines the Hausdorff distance between two sets $A = \{a_1, ..., a_p\}$ and $B = \{b_1, ..., b_q\}$ as follows:

$$H_W(A, B) = \frac{1}{p} \sum_{i=1}^{p} \rho_i \min_{b_j \in N_B^{a_i}} (Z_{i,j} \|a_i - b_i\|),$$

where $Z$ is the confidence matrix defined as the product of the difference of the eigen-vector for the Gaussian of the difference between each point in $A$ and $B$ [WC05a].

Sim et al. introduced the Least Trimmed Square Hausdorff Distance (LTS-HD) and defines the directed distance $h_{LTS}(A, B)$ as

$$H_{LTS}(A, B) = \frac{1}{S} \sum_{i=1}^{H} D_B(a_i),$$

where $S$ is the number of smallest distances to be used and $D_B(a_i)$ represents the $i$th distance in the sorted sequence $D_B(a_1) \le D_B(a_1) \le ... \le D_B(a_{N_A})$ [SKP99]. This distance is claimed to be more robust to occlusion, distortions and outliers [WC06a].

## 2.4 Transformation space

Algorithms differ in the kind of transformations that model the registration. Allow-able transformations may be rigid or non-rigid.

### 2.4.1 Rigid transformations

The search space is assumed to be the space of all allowed rigid transformations. A rigid transformation in $\mathbb{R}^2$ allows translations and rotations and may be written as

$$t_{rigid} = (t_x, t_y, \theta)$$

where $t_x$ and $t_y$ correspond to a translation along the $x$ and $y$ axes, and $\theta$ corresponds to a rotation. In other words, because a rigid transformation is defined by three parameters, the space of all the rigid transformations is $\mathbb{R}^3$. The search space for rigid transformations can be bounded since the input point sets are finite.

### 2.4.2 Affine transformations

Affine transformations, like rigid transformations, are linear transformations of the space. An affine transformation in $\mathbb{R}^2$ can be written in the form $p' = Ap + B$ . More precisely, if the point $p = (p_x, p_y)$ and $p' = (p'_x, p'_y)$, then an affine transformation can be written as follows:

$$\begin{bmatrix} p'_x \\ p'_y \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} p_x \\ p_y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} .$$

Since the four entries in matrix $A$ and the two in vector $B$ define an affine transformation, the search space of all affine transformations is $\mathbb{R}^6$. Once again, this space can be bounded using the finite input point sets.

### 2.4.3 Non-rigid transformations

**Optical Flow**

The concept of optical flow is used in the field of computer vision in order to recover the relative motion of an object and the viewer between two successive frames of a

temporal sequence. The fundamental assumption is that the brightness of a point stays constant, i.e.,

$$I(x, y, z, t) = I(x + \delta x, y + \delta y, z + \delta z, t + \delta t).$$

Using a Taylor expansion and ignoring higher-order terms, we get

$$\frac{\partial I}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial I}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial I}{\partial z}\frac{\partial z}{\partial t} + \frac{\partial I}{\partial t} = 0.$$

which can be written in vector form as follows:

$$\triangle I + \nabla I \cdot \mathbf{u} = 0,$$

where $\mathbf{u}$ describes the motion between the two images.

The optical flow determines a dense pixel-wise displacement field and is used for non-rigid registration but it does not guarantee the preservation of topology and coherence of a shape after deformation (two contiguous points can diverge to distant locations, disconnecting the shape and creating gaps). To alleviate this problem, regularization and smoothness constraints can be included in the model. Still, this problem makes optical flow unfit for shape registration and much less for extension to non-rigid registration of medical images.

**Thin Plate Spline**

Thin Plate Spline (TPS) is perhaps the most popular point-based non-rigid registration strategy. TPS minimizes the energy of the functional

$$I_f = \iint f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2 \, dx \, dy,$$

where $f(x, y)$ can be expressed as follows:

$$f(x, y) = a_1 + a_x x + a_y y + \sum_{i=1}^{n} w_i U \left( \| (x_i, y_i) - (x, y) \| \right),$$

and $U(r) = r^2 \log r$. To guarantee continuity it is required that $\sum_{i=1}^{n} w_i = 0$ and $\sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i y_i = 0$. The energy functional is minimized over the 2D plane and therefore the double integral for $I_f$ is computed over $\mathbb{R}^2$.

Given that $f(x_i, y_i) = v_i$ the coefficients $w_i$ can be found by solving the system

$$\begin{pmatrix} K & | & P \\ - & + & - \\ P^T & | & 0 \end{pmatrix} \begin{pmatrix} w \\ - \\ a \end{pmatrix} = \begin{pmatrix} v \\ - \\ 0 \end{pmatrix},$$

where

$$K = \begin{bmatrix} 0 & U(r_{12}) & \cdots & U(r_{1n}) \\ U(r_{12}) & 0 & \cdots & U(r_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ U(r_{n1}) & U(r_{n2}) & \cdots & 0 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix}.$$

To relax the interpolation, a regularization term can be added to have the function to minimize

$$H[f] = \sum_{i=1}^{n} (v_i - f(x_i, y_i))^2 + \lambda I_f$$

where $\lambda$ is the regularization parameter and controls the smoothing of the interpolation.

TPS receives pairs of points as input and generates a smooth surface that connects the points. This is done using a closed form making it fast. However, the problem is the lack of robustness towards wrong correspondences. The model cannot recover from misassignments. Therefore the main concern when using TPS is the automatic assignment of correspondences between pairs of points.

**Other deformable models**

Kass et al. present a spline-based model where the 2D contour detection problem in intensity images is the minimization of the energy functional $E_{snake} = \int E_{int}(v(s)) + \int E_{image}(v(s)) + \int E_{con}(v(s))$, where $E_{int}$ represents the internal continuity of the spline and can be tuned to allow corners [KWT87]. $E_{image}$ is a functional based on the image itself and can be a functional of pixel intensity, gradient or zero-crossings. $E_{con}$ models external forces to be performed by a higher level system or a user in an interactive system. Terzopoulus presented a symmetry-seeking 3D shape model based on a rolled sheet of elastic material and a deformable spline coupling forces which keep symmetry opposing extrinsic forces which force to keep the shape consistent with 2D projections [TWK88]. Pizer et al. present a model that represents shapes as a hierarchical decomposition of figures described by medial primitives and a probability model [PFY$^+$99]. The authors describe the task of registration as finding the object representation of the reference image $R$ and the object representation of the target as $T(R)$ and $T$ being the desired registration transformation. Huang presents a global-local registration scheme that represent shapes implicitly using a distance transform and computes a global affine registration using mutual information followed by a local deformable registration performed using bsplines and minimizing the sum of squared differences of intensities of the distance transforms plus a regularization functional [HPM06].

## 2.5   Search strategy

In this section we focus on search strategies for point-based registration. We start with the closed form solution for marker based registration, which is the only current application of point-based registration accepted as a standard. Then, we describe

the ICP method which is the most popular of these methods. We continue with the branch-and-bound and hill climbing methods that are widely used in this work. We conclude with a method that clusters voxel intensities and manipulates them in a point-based fashion.

## 2.5.1 Marker-based registration

Accurate image registration can be performed effectively if some reference markers are available in the images. However, where natural markers are not available, artificial ones need to be introduced. Since skin is motile, markers must be placed onto rigid anatomical sites through a surgical procedure, making this a potentially invasive procedure, especially if high precision is sought [Mur02]. This makes markers undesirable.

Registration of markers can be achieved by finding the least-squares estimate for the norm of the difference between two sets of points, which can be solved using a *singular value decomposition* (SVD)-based method. It is worth mentioning that this method requires that both sets of points have the same size, it is not robust to missing or noisy points and the correspondence between the point sets has to be explicitly provided. Establishing this correspondence requires an *a priori* approximate solution of the registration.

Let $A$ and $B$ represent the reference and target coordinates of the markers in 3D. Thus, both $A$ and $B$ can be represented by two $n \times 3$ matrices, where $|A| = |B| = n$. The problem is to find the least-squares estimate for $T$ and $R$ that minimizes the following error function:

$$\varepsilon^2 = \|A - BR - T\|.$$

The SVD approach was first presented by Schonemann [Sch66]. Umeyama [Ume91] and Kanatani [Kan94] introduced several improvements. This approach involves the following three steps.

1. The point sets $A$ and $B$ are translated in such a way that their mean is the origin. This defines the translation vector, $T$.

2. Compute the decomposition $UWV^T = B^T A$.

3. The rotation matrix $R$ is defined as follows:

$$R = U \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & det\left(UV^T\right) \end{pmatrix} V^T,$$

where $U$ and $V$ are the eigenvalue and eigenvector matrices computed by the SVD.

We use this method to compute our ground truth. It requires the determination of point correspondences. It is not robust in the sense that even if only one marker location is not carefully determined, the result is adversely affected.

## 2.5.2   ICP method

The *Iterative Closest Point* algorithm (*ICP*) is one of the most popular point-based registration methods. In each iteration, the algorithm selects a correspondence between the points in the two sets and calculates the rigid transformation that minimizes the squared sum of the residuals. It was first described by Besl and McKay [BM92] and several variants have been presented since then. One such variants is the *Trimmed ICP* (*TRICP*) (used in this dissertation for evaluation purposes) and adds a robustness quantile $q$ to the ICP optimization [CSK05] in a manner similar to the robust Hausdorff distance presented in our work. The TRICP algorithm is described as follows:

1. For each point $a_i$ in $A$, find its closest point $b_i$ in $B$ and compute the square of the distance $d_i$ between them, i.e., $d_i^2$.

2. Sort $d_i^2$ and select the $k = q |A|$ smallest values and calculate their sum (sum of *trimmed squares*) $S_{TS}$.

3. If the maximum number of iterations is reached or if $S_{TS}$ or the change in $S_{TS}$ are too small, stop. Otherwise, let $S'_{TS} = S_{TS}$ and proceed.

4. Compute for the $k$ pairs the rigid transformation $(R, t)$ that minimizes $S_{TS}$.

5. Transform $A$ accordingly and go back to step 1.

Four algorithms to perform the minimization for step 4 were previously compared [ELF97], including the popular SVD method [AHB87]. Eggert suggests that the method by which one computes the eigensystem of a matrix of the unit quaternion representation of the rotation [Hor87] is the most robust to noise and degenerate point set distributions, and is relatively fast.

The main disadvantage of the ICP is its sensitivity to start conditions. The method fails totally when the convex hull of the point sets do not intersect. ICP can be used for rigid and non-rigid registration. In rigid registrations performs acceptably but it is not robust to slight warping in the images. For non-rigid registration, it is very sensitive to starting condition and the deformable models converge easily to local minima.

### 2.5.3 Branch-and-bound

The geometric *branch-and-bound* framework was first developed by Huttenlocher et al. [HR93] and was used to perform approximate point pattern matching. A similar strategy was also proposed and implemented by Mount et al. [MNL98]

for the purpose of registering satellite images from which point features had been extracted.

The branch-and-bound search strategy can be used if rigid or affine transformations are allowed. It divides the (bounded) transformation space into non-overlapping hyperrectangles called *cells*. The *branching* strategy involves subdividing the cells into subcells, while the *bounding* strategy at any stage of the search involves deciding which of the cells need not be searched for a better transformation. Note that a cell contains an infinite number of possible transformations. Our algorithm only considers representative points (corners, center, or random representatives) within each cell. Each cell can be described in terms of a lower left and an upper right corner, which corresponds to the minimum and maximum coordinates in both dimensions. The transformations corresponding to these two corner points are denoted by $t_{lo}$ and $t_{hi}$.

The *uncertainty region* associated with a given point $a \in A$ and a given cell $T$ is defined as a rectangle (in the space of input points) that bounds all the possible mappings of $a$ under any transformation in $T$. This rectangle is defined by the transformed points $t_{lo}(a)$ and $t_{hi}(a)$ that lie at the corners of the uncertainty region. Uncertainty regions help to find the best similarity in a cell in an optimistic way, thus helping to define a lower bound of the Hausdorff distance for a particular cell, which in turn is used by the bounding strategy. Given a point set $A$, the *size* of a cell (set of transformations) is defined as the largest of the longest sides of the uncertainty regions (rectangle in the input point space) associated with all the points in $A$ and the cell. Given a collection of unexplored cells, the size of a cell may be used to order the search.

The branch-and-bound technique maintains the best solution found so far, i.e., the transformation, $t_{best}$, with the least weighted partial Hausdorff distance is main-

tained, along with the corresponding distance $D_{best}$. The search strategy maintains a priority queue of unexplored and promising cells. The queue starts with 1 cell, that corresponds to the bounds of the search space for the given input point sets. The priority value is the "size" of the cell.

To explore a cell $T$, a lower bound $D(T)_{lo}$ (minimum Hausdorff distance that can be achieved using the transformations in this cell) is first calculated. If the lower bound is at most equal to the best similarity found so far $(D_{best})$, then this cell is incapable of improving $D_{best}$ and is discarded. Thus the cell $T$ is "killed" if $D(T)_{lo} \geq D_{best}$. For a cell $T$, the lower bound is computed as the Hausdorff distance between the point set $B$ and the uncertainty regions associated with points in $A$ and cell $T$. The distance between a point $P$ and a rectangle $R$ is defined as the minimum distance from $P$ to any point on $R$ (0 if $P$ is contained in $R$). If a cell is not killed, the Hausdorff distance for the transformation corresponding to the center of the cell is used as a "witness" upper bound (denoted by $D(T)_{hi}$). If $D_{best} > D(T)_{hi}$, then $D_{best}$ is updated. For the branching step, a cell is split in half along the direction that most contributes to the size of the uncertainty region producing two new cells that are inserted back in the priority queue. In summary, given a bounded space $T$ of transformations and two point sets $A$ and $B$, the problem of point matching can be proposed as finding the transformation $t_{best} \in T$ that minimizes the distance between $A$ and $T_{best}(B)$.

Mount et al. showed that the algorithm can be speeded up by introducing three different types of relaxations to the rules of the algorithm [MNL98]. These relaxations are controlled by three error parameters that may be controlled by the user. Firstly, the algorithm may be stopped if the Hausdorff distance for the best transformation can be guaranteed to be within $\epsilon_a$ of the optimal alignment. Secondly, the quantile $q$ used in the partial Hausdorff distance is allowed to be imprecise up

to a factor of $1 - \epsilon_q$. Thirdly, a cell may be killed as long as it cannot improve the best transformation by more than a factor of $1 + \epsilon_r$. Using the above criteria, a cell $T$ is killed if

$$D(T)_{lo} \geq \frac{D_{best}}{1 + \epsilon_r}, \text{ or } D(T)_{lo} > D_{best} - \epsilon_a.$$

The algorithm stops when all the cells are killed or when $D_{best} < \epsilon_a$.

We used this approach because it is robust and explores the space in an orderly fashion. The main challenge in using a branch-and-bound approach resides in the computation of the lower bound for a group of transformations which limits its applicability to rigid and affine transformations. Often the lower bounds computed were low enough that the use of approximate starting points was useless, and it does not provide sufficient "bounding".

## 2.5.4 Stochastic Hill Climbing

Stochastic hill climbing is a heuristic for solving optimization problems and can be considered as a generalization of genetic algorithms [KPP95, RK96]. In any given iteration of the algorithm, it generates a randomly sampled population $Q$ of solutions sampled from a normal distribution $\overrightarrow{N}(\overrightarrow{\mu}, \overrightarrow{\sigma}^2)$ around the representative of the best points in the previous iteration. The vector $\overrightarrow{\mu}$ is initially set to the center of a $d$-dimensional bounded space and updated using Hebbian learning as explained below. The variance of the distribution $\overrightarrow{\sigma}^2$ is initially chosen in some appropriate manner and then decreased in each iteration. The subset $Q_{best}$ of the best vectors in this set, in terms of minimizing Hausdorff distance, is used to shift $\overrightarrow{\mu}$ in the direction of its mean, $\overrightarrow{\mu}_{best}$. Hebbian learning is used to guide the update of $\overrightarrow{\mu}$, and is given by

$$\overrightarrow{\mu} = \overrightarrow{\mu} + L\left(\overrightarrow{\mu}_{best} - \overrightarrow{\mu}\right),$$

where $L$ corresponds to the learning coefficient. The larger the value of $L$, the more drastic is the shift toward the best population, and higher is the probability of getting trapped in a local minimum. In practice, this approach converges very quickly to the vicinity of the globally optimum solution, although the theoretical convergence cannot be guaranteed.

The search space corresponds to $\mathbb{R}^d$, where $d$ is the number of dimensions of the transformation space. Note that for 2D images embedded in a lattice $P$, $d$ corresponds to two times the total number of control points which is equal to $2NM$.

Given two point sets $A$ and $B$ and the parameter $k$, the image registration problem can be thought of as minimizing the following function:

$$\Gamma_{H_k} : \mathbb{R}^d \to \mathbb{R},$$

where $H_{OPTIM}$ is the robust partial Hausdorff distance (See Eq. (2.1)). The function assigns a real value to a transformation $t$ described in terms of its $d$ parameters:

$$\Gamma_{H_k}(t) = H_{OPTIM}(A, t(B)).$$

The image registration problem is that of finding the transformation $t_{best}$ with parameters represented as a vector $\overrightarrow{v}_{best}$, where

$$\overrightarrow{v}_{best} = \arg\min \Gamma_{H_{OPTIM}}(\overrightarrow{v}), \quad \overrightarrow{v} \in \mathbb{R}^d.$$

The transformations, $t_{lo}$ and $t_{hi}$, defining the corners of a cell, can be used to create an initial distribution $\overrightarrow{N}_0(\overrightarrow{\mu_0}, \overrightarrow{\sigma_0}^2)$.

We found that this heuristic converged quickly to the neighborhood of the global solution but was unable to find the optimal solution. The main drawback of this approach is its probabilistic nature and the lack of computational bounds, convergence and repeatbility. In our experiments we found that the variances for repeated experiments were generally low. In the non-rigid case it worked well with the multi-resolution free form deformation model.

## 2.5.5 Shape Context

Belongi et al. introduced a descriptor called *shape context*, which is a histogram of the locations of all the points in the point set, relative to a reference point [BMP01]. The authors use a log-polar coordinate system and 12 bins for the rotation and 5 bins for the log of the radius. The cost of matching two points $p_i$ and $q_i$ is $C(p_i, q_i) = C_{i,j} = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]2}{h_i(k) + h_j(k)} C_{ij}$. The bipartite problem is solved using a method that find a permutation $\pi(i)$ that minimizes $\sum_i C_{i,\pi(i)}$. $C$ is a square matrix. To deal with different cardinalities or to add robustness to outliers, dummy nodes con be added with a constant matching cost of $\epsilon_d$. Once the correspondences have been established, a regularized TPS is used to model the non-rigid transformation. The matching algorithm iterates these two steps until a fixed number of iterations is reached. The two step method is very similar to ICP.

This method assigns a signature to each point and computes a cost for matching pairs of points. Similar costs may lead to wrong correspondences. The size of the cost matrix $C$ is square in the size of the point sets and solving the bipartite problem for large point sets can become computationally expensive. The main problem of the approach presented by Belongi is related to the high sensitivity of the TPS deformable model to wrong correspondences.

## 2.5.6 HAMMER

HAMMER is a non-rigid registration method that uses tissue classification data to select driving points for the registration [SD02]. Therefore, it requires previous segmentation of the images. HAMMER uses brain MRI images with three marked types of tissues: white matter, gray matter and other. The method assigns a vector to each point (voxel) in the image. This vector contains information such as the type

of edge, the intensity at the voxel and invariant moments. The latter is computed for each type of tissue, giving priority to boundary points to drive the registration. An energy function is presented to compute the cost of drawing a correspondence between two points. This function is not only of the particular points but also their neighborhood. The terms of this function include the cost of mapping the reference to the target, the cost of mapping the target to the reference and a smoothness. The authors have also included several heuristics for brain tissue registration. This is a very interesting method that takes advantage of the segmented tissue data. This data is not available for general registration. The author mention wavelet and Gabor vectors as also suitable for registration. The advantage of the later is that the scale and resolution can be modified to compute different responses and not just the size of the neighborhood as in invariant moments.

HAMMER computes a cost function of assigning a correspondence between two points. Initially this function is evaluated in a small subset of voxels (driving points) that is iteratively enlarged until all the voxels in the image are included. The warping function is not clearly defined but the term local affine transformations is used throughout the article. The error of 0.966 mm is related to the average measured displacement error with a maximum of 1.875.

A hierarchical method is used to evaluate only points with large differences in their vector representations and iteratively include more and more points. It allows many-to-many correspondence assignments but an entropy factor systematically reduces the probability of these happening until you have only one-to-one assignments for later iterations. This provides robustness to the method. The final deformation field is smoothed using the Laplacian operator on a neighborhood around each point.

## 2.6   Error measurement

To measure the error we tried first to measure the error in the rigid transformation space, this is measuring the absolute value of the difference of each rigid parameter (translation in $x$, rotation) from the ground truth which was always computed using SVD. This method, although very good to analyze how a particular method is exploring the search space, has several disadvantages. First, the experiment accuracy cannot be compared with other methods who use different transformation spaces (i.e., non-rigid), second, it assigns an accuracy value to rotations and translation, which can not be compared, particularly because rotations of points far away from the origin correspond to large translations in space. And third, the error in computing the marker registration is propagated into the experiment accuracy evaluation; i.e., there are measurement errors in the marker localization (the residuals of the SVD are not zero) and any direct comparison to this SVD result will add that uncertainty into the results.

We used the method proposed by Maurer et al. [MFW$^+$97] to achieve a more general and marker-independent accuracy indicator known as *target registration error* (TRE). The method is composed of the following steps:

1. Compute the marker-based registration $S$ with FLE > 0.

2. Select an arbitrary set of markers $U$ around the target area of interest and compute their ground truth mapping $V = S(U)$.

3. Compute the rigid transformation $T$ that registers the point sets.

4. Compute TRE for each target as $\|v_k - T(u_k)\|$.

The TRE is a more objective measure of accuracy of registration because it computes the distance between corresponding points other than those used to estimate the ground truth.

CHAPTER 3

## POIRE: A METHOD FOR POINT-BASED REGISTRATION

Exploring point-based registration methods and making them more accurate than existing methods is one of the main goals of this dissertation. In order to achieve this goal we explore the registration of images in radiotherapy. Image registration in radiotherapy is challenging and interesting because of the characteristics of the images acquired at high energy. As mentioned in Section 2.1.4, portal images are characterized by low quality and high scatter. In this chapter we will describe in detail the process of patient position verification and also describe the preliminary experiments performed to evaluate the suitability and accuracy of point-based registration in radiotherapy for simulated *EPID* and *DRR* images. After these preliminary experiments, we describe PoIRe, our registration framework. We study PoIRe in three different settings. First, PoIRe is used to register pairs of EPID images of an *anthropomorphic cranial phantom*. Second, we use it for CT-CT registration. Third, we analyze PoIRe for multi-modal registration of CT-MR real patient data using benchmark data from the *Retrospective Image Registration Evaluation* (RIRE) project.

## 3.1 Introduction

This chapter presents PoIRe, a novel framework for robust point-based image registration of 2D and 3D medical images based on the partial Hausdorff similarity measure. PoIRe uses two well known search strategies and develops a new hybrid search strategy. We use the robust, partial Hausdorff distance as our dissimilarity measure.

Image registration in radiotherapy has extensive applications. Mainly, it is used to (1) integrate information from different modalities for tumor and critical anatom-

ical structure identification in treatment planning; and (2) verify patient position with respect to coordinate system of the medical linear accelerator (or LINAC) used for radiation beam delivery as part of treatment verification. Treatment outcome is intrinsically linked to precise beam delivery and patient positioning, which are needed to maximize radiation dose to the tumor and minimize its effects on neighboring critical and normal anatomical structures. Typically, verification of patient position is carried out by weekly portal imaging prior to radiotherapy treatment using megavoltage x-rays, which are of the same energy used for treatment in radiotherapy. There are several imaging modalities involved in treatment verification. Patient position can be verified either by comparing successive portal images, or by comparing the portal images with the digitally reconstructed radiographs (DRRs) reconstructed from axial computed tomography (CT) imaging. Portal images have been traditionally acquired using film but with improving digital technology, better quality portal images can be acquired using an electronic portal imaging device (EPID). Computed tomography (CT) volumes are used during diagnosis and also in patient treatment planning calculations. These 3D kilovoltage X-ray images display good contrast and high resolution. CT images are needed for both anatomical target identification and for modeling radiation beam interactions. As described in Section 2.1.1, digitally reconstructed radiographs (DRRs) are 2D projections of CT images for a specified source-patient-detector geometry. The CT (and DRR) provides the "reference" images regarding the intended patient position in the sense that the beam planning assumes a specific alignment of the patient with the LINAC coordinate system. The process of aligning the patient on the treatment couch with the CT volume is called patient verification. Typically, verification of patient position is carried out by 2D portal imaging immediately prior to radiotherapy treatment using megavoltage X-rays. Patient position can be verified either by comparing suc-

cessive portal images, or by comparing the portal images with previously captured DRRs. Although 2D portal imaging is most common, 3D Conebeam CT (CBCT) is also becoming increasingly popular [MMA06]. CBCT is used for high precision patient positioning achieved by registering CT and CBCT imaging. As mentioned in chapter 2, MRI is now established as an important complement to CT imaging for treatment planning.

The above mentioned properties of EPID images makes their registration a serious problem. Voxel-based methods are challenged because of the high noise levels: Voxel-based similarity functions such as sum of squared differences (SSD) and cross-correlation (CC) are too sensitive to the non-Gaussian noise present in the images. Mutual information (MI) is a more promising choice and we compare our results to methods that rely on this similarity measure.

PoIRe is a feature-based method in which points of interest are extracted using classic edge detection methods. The high level of scatter in EPID images is reflected in the presence of artifacts in the extracted point sets. Therefore, we need to use a similarity measure that is robust to these artifacts. We use the partial Hausdorff distance. The properties of edges extracted from EPID images pose a problem for the ICP method and its variants. These methods are not robust to the high level of scatter points included in the images.

We show in our preliminary experiments with simulated images that the point-based schema presented produces accurate registrations. The evaluation of these experiments is possible because we know the registration results beforehand. For real images, it is difficult to evaluate this accuracy. For this purpose we use retrospective registration which plays an important role in the evaluation of registration methods [WFW+97]. As part of our evaluation process we use the *Retrospective Image Registration Evaluation* (RIRE) project created by Fitzpatrick from Vanderbilt

University to provide an objective web-based method to evaluate the accuracy of rigid registration methods for brain images in real patients.

The RIRE project has available data for CT, MR and PET images. MR images were corrected for static field inhomogeneity and scale distortion. Several steps were designed to compute the ground truth for every pair of images. The first step was fiducial marker localization. Next the fiducial-based registration was performed as described in Section 2.5.1. The fiducial registration error (FRE) is defined as the root mean square (rms) of the distance between corresponding markers. Finally, the fiducial markers and stereotactic frames were removed from each image. The RIRE website provides user name and password to participants and report all validations performed of the system. The correct answers from the fiducial-based registration are kept secret. To validate a registration, the user has to provide the mapping of the 8 corner points of the volume to register and the system outputs the TRE for 10 areas of interest in the brain. This process makes it impossible to guess the right answers from successive testings on the system.

The method described in this chapter, PoIRe, incorporates three distinct search strategies. First, it uses a *branch-and-bound* search (Section 2.5.3), loosely based on the work of [MNL98]. The second strategy is a *stochastic hill climbing* search (Section 2.5.4). The third method, a newly proposed hybrid strategy, uses hill climbing to generate quick approximate start points and then employs the branch-and-bound strategy. In this chapter we compare the three strategies and study the accuracy and robustness of PoIRe to noise; we also study the effect of sampling on extracted point sets.

For PoIRe, the search space is the space of all transformations. Both *rigid* and *affine* transformations have been previously considered for medical image registrations. Rigid transformations allow for translations and rotations, while affine

transformations allow additionally for scaling and skewing. In our experiments, we only searched the space of rigid transformations, although PoIRe can be easily modified to handle affine transformations as well. It has been shown that rigid transformations are sufficient for accurate cranial image registration [FWM98]. We parametrized the rotation and translation explicitly. For 2D images, we represent rigid transformations as a vector of three components (one for rotation and two for translations). For 3D images, they are represented as a vector of six values (three for rotations and three for translations).

PoIRe specifically avoids assumptions about the availability of markers and is designed to perform well in their absence. Our experiments perform intramodal, point-based image registration of 2D EPID images, which are used to study variations in patient positions between and during radiation treatments. We also study the much harder problem of intramodal, point-based registration of 3D CT images.

In all the experiments, the reported errors are relative to the marker-based registration used as a gold standard. We discuss experiments with a range of parameters and discuss how they should be chosen. Our 2D imaging data has applications to patient positioning in cranial radiotherapy.

Although this chapter focuses on EPID-EPID and CT-MR registrations, in the following section we will describe in detail the preliminary experiments results for registration of simulated images for EPID-EPID and EPID-DRR. This preliminary work was our foundation to determine that point-based registration of EPID images had a potential for accurate registration and motivated the further development of PoIRe. This preliminary work also showed the limitations of point-based registration for EPID-DRR and we did not pursue that direction any further (in Chapter 4 we will use local descriptors to solve this problem). The limitation of using gradient information as local descriptor was also reported.

## 3.2 Preliminary experiments

This section describes the experiments performed on EPID-EPID and EPID-DRR. This was our first attempt to use point-based registration for medical images. From this preliminary results, several improvements were motivated and their implementation came to be PoIRe. This section outlines the accuracy of point-based methods for intra-modal registration and also outlines the limitations for point-based multi-modal and multi-resolution image registration.

### 3.2.1 Simulated data sets

The preliminary experiments were performed on simulated data. The simulated data was designed in a way that would facilitate measuring the accuracy of our registration algorithms. Simulated images, consisting of an actual EPID image rotated and translated to known values using software, were used since position shifts could be exactly determined (limited only by pixelization effects). The experiments on simulated data were performed by generating the data as follows. Given a reference image $A$ and a random rigid transformation $t$, the target image $B$ was produced by applying $t$ to $A$, i.e. $B = t(A)$. Then, the feature extractor was used on $A$ and $B$ to produce two point sets $P_A$ and $P_B$ respectively. These point sets each had about 1400 points. Figure 3.1 shows the data generation step followed by the registration step.

The point sets $P_A$ and $P_B$ described were used as input to the point-based registration algorithm which produced an output registration $t_P$. The differences between $t$ and $t_P$ ($\Delta\theta$, $\Delta x$ and $\Delta y$) were recorded. For robustness, the quantile $q$ was set to 0.7, meaning that the algorithm was only required to match 70% of the points. The mean of the time and the errors were recorded. The last experiment
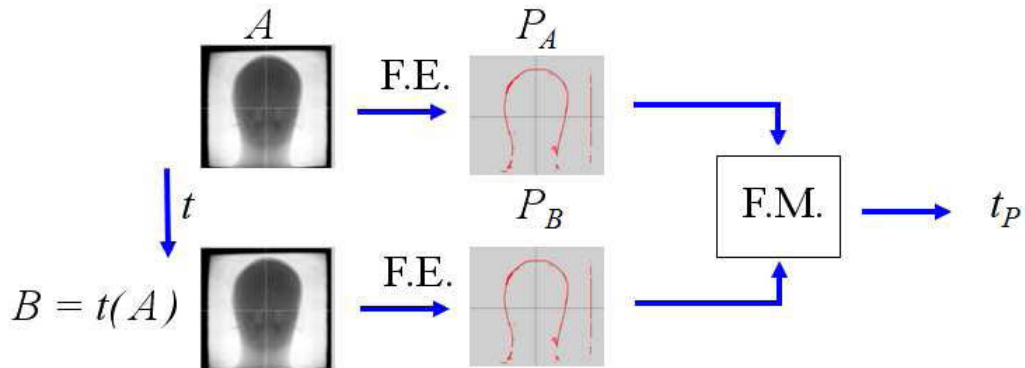
Figure 3.1: Generation of simulated data

for simulated data was performed with multimodality images using an algorithm that included all the best features as concluded by the first four experiments. Every experiment was repeated 1000 times. The computations woth simulated data were done with a 2.8GHz Intel processor with 2GB RAM using Red Hat Linux 3.4.4-2 operating system (1.4GHz Athlon processor with 256 MB RAM for simulated data experiments).

### 3.2.2 Preliminary results

The following results were obtained from the experiments using the simulated data sets. The point-based multimodality image registration algorithm presented here displays tremendous speed and accuracy; the accuracy exceeds the positioning accuracy of the patient on the treatment linac. To generate quantitative results we used DRR images, simulated EPID images, and actual EPID images. Simulated images, consisting of an actual EPID image shifted to known values using software, were used since position shifts could be exactly determined (limited by only pixelization effects). Actual EPID images typically have inherent setup errors since the imaging devices used to generate the CT data set and EPID images are differ-
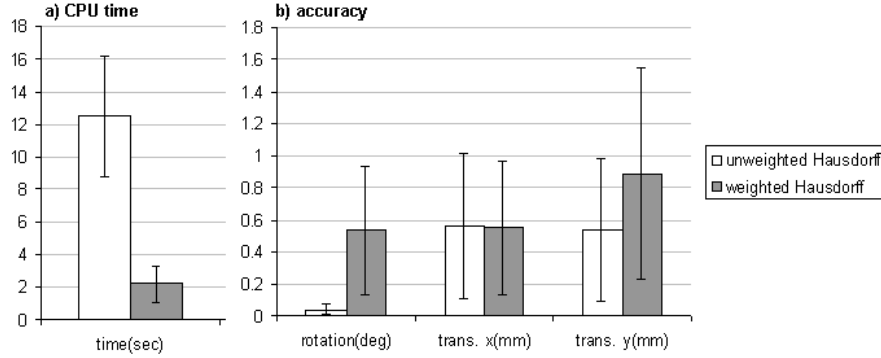
Figure 3.2: Comparison results for unweighted (Equation 2.1) and weighted (Equation 2.1) Hausdorff distance using the branch-and-bound intramodal image registration algorithm. Both reference and target images were EPID phantom images. The number of points in each image was around 1400. 1000 registrations were performed. Graphs show a) average CPU time for each registration, and b) mean errors in the rotation, translation in x, and translation in y for the computed registrations.

ent. The actual EPID images were used to study the effect of imaging noise on the registration.

**Weighted Hausdorff Distance**

The first set of experiments was set up to evaluate the improvement in performance by comparing two implementations of the same registration algorithm, one using an unweighted Hausdorff distance and another using the weighted Hausdorff distance. As shown in Fig. 3.2, using the weighted Hausdorff distance, the average computation time was reduced by 82% over the unweighted case. The significant improvement occured because the weighted metric favored strong edges, and thus drastically reduced the number of cells inspected. Even though the average and worst case errors were much higher for the weighted measure, the errors were within acceptable bounds for our application.
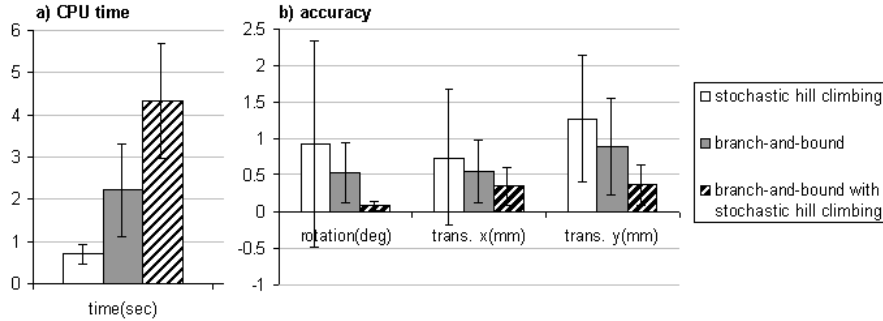
Figure 3.3: Comparing Search Strategies. Three different search strategies were compared: pure stochastic hill climbing, branch-and-bound, and branch-and-bound with stochastic hill climbing approach. All of them used the weighted Hausdorff measure for intramodal image registration. Both reference and target images were EPID phantom images. The number of points in each image was around 1400. 1000 registrations were performed. Graphs show: a) mean CPU time for each registration, and b) mean errors in the rotation, translation in x and translation in y for the computed registrations.

**Stochastic Hill Climbing**

The second set of experiments shown in Fig. 3.3 explored the weighted Hausdorff distance in three different scenarios: a pure stochastic hill climbing strategy, the pure branch-and-bound method and finally, the combined strategy of branch-and-bound with stochastic hill climbing. The combined strategy of branch-and-bound with stochastic hill climbing resulted in a significant reduction of mean error (84% for rotation, 35% for translation in $x$ and 59% for translations in $y$), but at a two-fold increase in computation time over the branch-and-bound, and a four-fold increase over the stochastic hill-climbing strategy (Fig. 3.3). The pure stochastic hill climbing strategy was 84% faster than the strategy of branch-and-bound with stochastic hill climbing. However, the resulting worst-case translation errors was more than 2 pixels and the worst-case rotation error was more than 2 deg, both of which are unacceptable for our application.
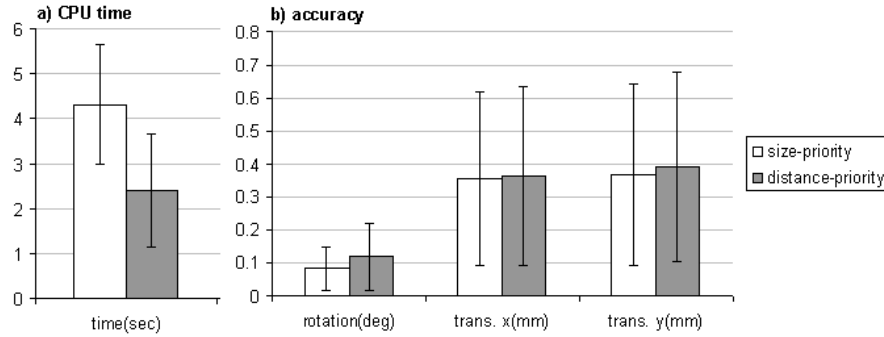
Figure 3.4: Comparing Size Priority vs. Distance Priority in the Branch-and-Bound Algorithm with Stochastic Hill Climbing Two different priority criteria for the processing of cells in the branch-and-bound strategy were compared. Both reference and target images were EPID phantom images. The number of points in each image was around 1400. 1000 registrations were performed. Graphs show a) mean CPU time for each registration, and b) mean errors in the rotation, translation in x and translation in y for the computed registrations.

**Distance-Based Priority**

In Fig. 3.4, an algorithm that prioritizes cells according to size was compared with a variant algorithm using distance-based priority. The distance-based priority turned out to be far superior to the one using size as the priority. As shown in Fig. 3.4, there is a speedup of a factor of almost 2 with a negligible increase in the average error.

**Multi-resolution Feature Extraction**

Fig. 3.5 shows a series of experiments with different levels of resolution used for feature extraction. Level 3 images, which contain eight times less information along each axis than level 0 images, can be registered more than 95% faster than the same images at level 0. While the accuracy in translations is almost the same at all levels, the error in the rotation angle increases from being negligible at level 0 to about 1.5 deg at level 3, a level that is not acceptable for clinical operations (Fig. 3.5). Even
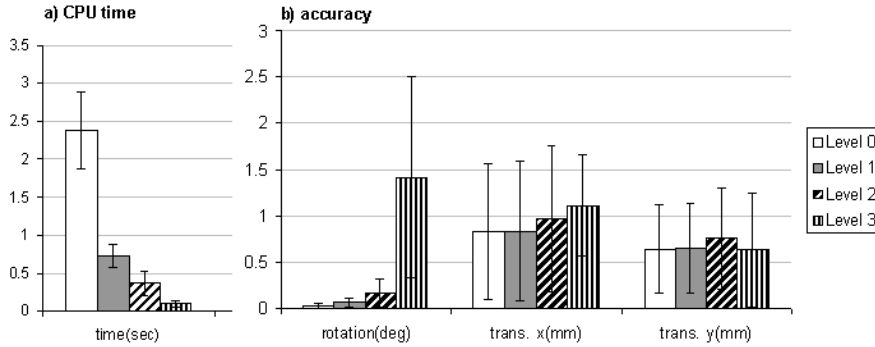
Figure 3.5: Comparing different levels of resolution Four levels of multiresolution scaling were compared. The registration algorithm corresponds to the branch-and-bound combined with stochastic hill climbing and with the distance priority. The number of points in each image was around 1400. 1000 registrations were performed. Graphs show a) mean CPU time for each registration, b) mean errors in the rotation, translation in $x$ and translation in $y$ for the computed registrations.

though the results at level 3 were not very accurate, they can be used as starting points for matching at higher resolutions in order to reduce overall computation times.

**Multimodality Image Registration**

A final experiment was designed to calculate the computation times and accuracy with images of different modalities using the combined strategy: branch-and-bound and stochastic hill climbing with also the distance based priority. The results are shown in Fig. 3.6. A DRR image was used as reference and simulated EPID images were used to measure the accuracy of the image registration. Fig. 3.6 shows very acceptable computation times below 1.5 seconds, error in rotations close to 0.5 deg and translation errors close to 1.5 pixels in the $x$ dimension and 1 pixel in the $y$ direction.
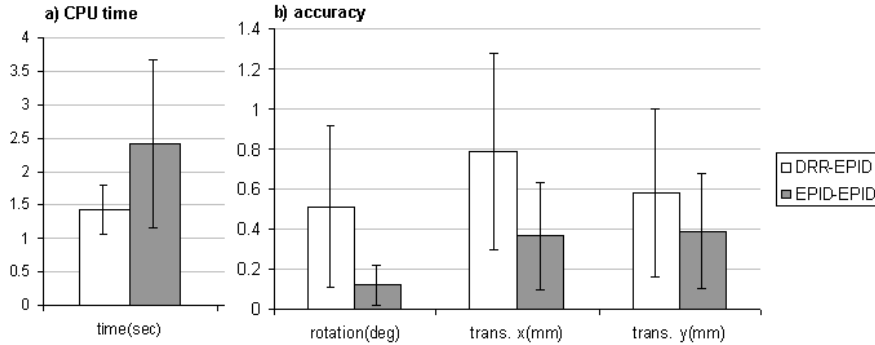
Figure 3.6: Intramodality vs. Multimodality Image Registration. EPID-EPID intramodality registration was compared with DRR-EPID multimodality registration. The registration algorithm corresponds to the branch-and-bound combined with stochastic hill climbing and with the distance priority. The number of points extracted was around 1400 and 1000 for the EPID and DRR images respectively. 1000 registrations were performed. Graphs show a) mean CPU time for each registration, b) mean errors in the rotation, translation in x and translation in y for the computed registrations.

**Convexity of Hausdorff distance**

It was of interest to evaluate the form of the Hausdorff distance for optimization. Figure 3.7 shows that the Hausdorff distance is monotonic (does not present local minima). We fixed one of the three rigid parameters and evaluated the Hausdorff distance at different locations of the plane for the two free parameters. Having a single global minimum, being smooth and convex for pair-wise dimensional evaluations are good properties for the Hausdorff distance.

### 3.2.3 Conclusions for preliminary experiments

The preliminary experiments showed that point-based registration of simulated images be accurate and suggested extending to the additional experiments shown in the next section. It showed large errors for multi-modality registration suggesting that a different approach for point-based registration needs to be used. In chapter 4
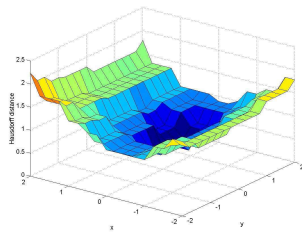
(a) translation in x and y     (b) translation in x and rotation     (c) translation in y and rotation
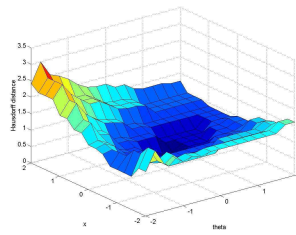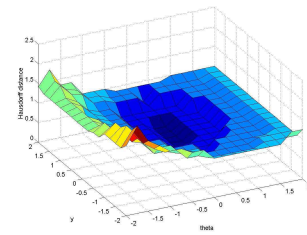
Figure 3.7: Convexity of Hausdorff distance

we will see how the errors get substantially reduced using local descriptor information. It showed the fast convergence of the hill-climbing method and the accuracy of running branch-and-bound followed by hill-climbing. This motivated the development of the Hybrid search strategy that will be shown later in this chapter to be robust and accurate.

## 3.3 Method

After several improvements from the preliminary set of experiments, we created the registration framework that is described in this chapter. In this section we first describe the feature extraction process used to convert images to point sets. Next, we describe the search strategy for robust and accurate registration, followed by a highlight of our improvements to the search strategy. We finish with a description of the standard error evaluation used in evaluating rigid registration.

### 3.3.1 Feature extraction

The feature points from an image consists of the set of points (2D or 3D) extracted from the original image (See Figures 3.8(b) and 3.8(c)). Feature points were ex-

tracted using classical edge detection algorithms. As suggested by our discussion in Section 2.2, we used Canny's algorithm [Can86] for EPID images and the Deriche algorithm [Der93] for CT and MR images. In all our experiments we automated the edge extraction methods, avoiding any user interaction to select important points. The extracted points have as origin the radiation center of the image.

### 3.3.2 Hybrid method

Borrowing ideas from a multi-resolution registration approach, we present a two-phase approach that allows achieving fast results while producing small variances in the accuracy results. In the hybrid approach we want to take advantage of the fast convergence of the stochastic hill climbing to a small neighborhood and the guaranteed convergence of the branch-and-bound.

**Phase 1**  We run the stochastic hill climbing together with very small sampling probabilities to compute fast and approximate registrations.

**Phase 2**  We use the best transformation found in phase 1 to start the branch-and-bound search. To reduce the running times while achieving the desired accuracy, we added one more innovation, i.e., we reduced the search space and centered it around the approximate answer. This was necessary because the branch-and-bound was unable to take advantage sometimes even when the starting point was a transformation close to the global optimum because, in general, the lower bounds are small and very few cells are killed based on the similarity of the starting point. To reduce the running times while achieving the desired accuracy, what we found more useful, was to reduce the search space and center it around the approximate answer.

(a) Markers
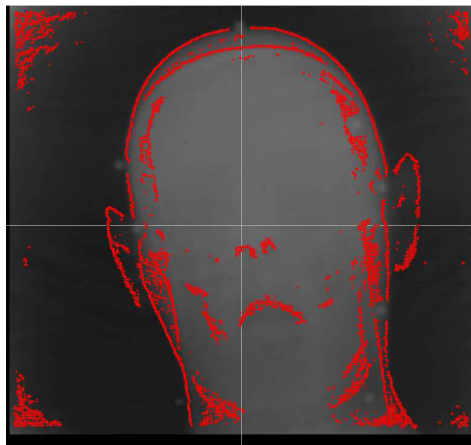


(b) Feature points for EPID
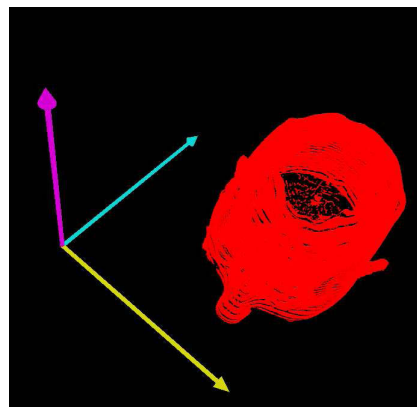


(c) Feature points for CT

Figure 3.8: Markers and feature points: (a) shows the 4 simulated marker positions in the reference image (shown as circles) that were projected into each one of the target images using the inverse "ground truth" transformation. These markers are used for computation of the TRE for all our EPID-EPID experiments. (b) shows an example of feature points extracted from an EPID image using the Canny algorithm. It can be seen that there are feature points associated with radiation scatter (see left bottom and left top corners) but were not removed. The only preprocessing of the point sets was the removal of the points around the fiducial markers to avoid biasing the point-based registration methods results using points corresponding to marker edges that are not generally available. (c) shows an example of feature points extracted from a CT image using the Deriche algorithm.

The use of a distance measure between point sets that is robust to the presence of outliers helps to decrease the effect of noise and to perform accurate registrations on feature point sets of different sizes. The partial Hausdorff distance measure (see Equation 2.1) provides the necessary robustness. In radiotherapy, patient positioning on the LINAC is carried out based on rigid transformations, even though the image registration may utilize a more general group of transformations (e.g., affine, non-rigid) since out-of-plane transformations may be involved (as in, portal imaging) and nonrigid transformations of patient anatomy may be detected.

### 3.3.3   Improvements to the search strategy

Here we present several improvements of the basic search strategy outlined above. These include an improved distance metric to guide the search, a choice of several possible search strategies, a multi-resolution strategy, a more efficient computation of the Hausdorff distance, an improved order of processing the cells by the branch-and-bound strategy, and an improved upper bound computation.

**Gradient Weighted Partial Hausdorff Distance**

The partial Hausdorff distance gives equal importance to all matched points. Our weighted version assigns weights corresponding to the magnitude of the gradient vector at the point in the source image. Note that stronger edges correspond to higher gradients, and in x-ray imaging, with the exception of tissue-air interfaces, higher edges typically correspond to bony anatomy. Generally, one has more confidence in using bony anatomy as features for determining patient position because it is nondeformable, and the major bones are generally rigid with respect to each other. Thus a bias is introduced towards using "stronger" points to compute the

Hausdorff distance. The mathematical formulation is as follows:

$$H(A, B) = \max_{a \in A}^{k} \min_{b \in B} w_a w_b dist(a, b) \qquad (3.1)$$

where

$$w_p = \nabla I_A(p) = \frac{\partial I_A}{\partial x}(p) + \frac{\partial I_A}{\partial y}(p)$$

is the weight of point $p$, and $I_A$ is the image from which the point set $A$ was extracted. As shown later, this weighted measure improves the efficiency considerably. For the rest of this chapter, for brevity, we will refer to the "gradient weighted partial Hausdorff distance" simply as "weighted Hausdorff distance".

We implemented two other variants of our image registration algorithm, one with pure stochastic hill climbing heuristic, implemented by incorporating the stochastic hill climbing into the branch-and-bound strategy. In the second variant, the stochastic hill climbing was used when computing the upper bound for a cell. The improved estimates of the upper bound for a cell helps to kill more cells in the "bounding" process of the branch-and-bound strategy.

**Multi-resolution Feature Extraction**

The multi-resolution feature extraction uses downsampled or lower resolution images to produce an approximate match that can be used as a starting point for the process with images at higher resolution levels. More importantly, the lower resolution matches produce an approximation good enough to reduce the search space at higher resolutions. The original image has the highest possible resolution (level 0). The image at level $i$ is sampled at twice the frequency of the image at the next lower resolution level (level $i + 1$), thus producing an image with half the information on both axes. These ideas dramatically reduce the computation times without deteriorating the accuracy of the results for the matching at the highest resolution.

**Miscellaneous Improvements**

Here we list other improvements incorporated into PoIRe. (a) The basic strategy uses only one representative point (its center) per cell. Sampling a small collection of points in the cell can improve the upper bound for each cell, allowing for more cells to be killed later on. (b) The first part of the Hausdorff distance computation corresponds to a nearest neighbor search. Since the Hausdorff distance is measured a large number of times, the optimization of this step is essential for speedup. After experimenting with several data structures (R-, R* trees, etc.), k-d trees were chosen for speed and stability of the code. (c) In the basic strategy, the algorithm starts with one cell. Since it is rare for cells to be "killed" in the initial stages, the first $L$ steps of the branch-and-bound are modified to do only "branching". This speeds up the program with no loss in accuracy. (d) In the basic strategy, cells are processed in the order of the decreasing size of their uncertainty regions. We implemented a priority criteria based on the distance from the center of the cell to the current best transformation found so far, with closer cells processed first. It can be considered as a greedy technique because areas that are closer to the more promising points are explored first.

### 3.3.4 Error Evaluation

There are several reasonable ways to compute registration errors. A few of these were used and are described below. Fiducial markers were used to compute the gold standard (also referred to as the ground truth or correct answer).

**TRE and FRE**

We used the *Target Registration Error (TRE)* proposed by [MFW$^+$97] to achieve an objective measure of accuracy of registration because it computes the distance between corresponding points different those used to estimate the ground truth. The steps to compute TRE are as follows:

1. Compute the registration $S$ of the fiducial markers with some residual error.

2. Select an arbitrary set of markers $U$ around the target area of interest and compute their ground truth mapping $V = S(U)$.

3. Compute the rigid transformation $T$ that registers the point sets.

4. Compute TRE for each target as $\|v_k - T(u_k)\|$.

In our experiments, we have four target points (as seen in Figure 3.8(a)) per image. We define the *maximum TRE* (and *mean TRE*) as the maximum (mean, respectively) TRE over all 4 targets and over all the images in the input set and over all repetitions performed.

When evaluating the accuracy of the ground truth, we follow the same procedure as in the previous section, but using the fiducial markers as target points. For this case, we refer to the corresponding TRE as the *Fiducial Registration Error (FRE)*.

**Computing errors in transformation space**

Computing the error in transformation space involves computing the translational and rotational errors. Toward this end, the gold standard (i.e., correct answer) for our experiments was computed using SVD marker-based registration of the fiducial markers attached to each one of the phantoms used for our experiments. The SVD produces a transformation comprised of a rotation matrix and a translation vector.

For 2-dimensional images the rotation matrix is converted into a single rotation value. In 3D, this rotation matrix is decomposed into three Euler angles. The reference consisted of the rigid transformation parameters: one rotation and two translation values (in 2D) or of three rotation and three translation values (in 3D). PoIRe searches the space of the rigid transformation parameters instead of the space of an N-dimensional matrix and vector.

In 2D, a transformation $T$ is defined as $T = (\theta, x, y)$. Let the SVD gold standard and the best transformation found by PoIRe be represented by $T_G = (\theta_G, x_G, y_G)$ and $T_B = (\theta_B, x_B, y_B)$, respectively. The error in $T_G$ is represented by the transformation $T_e = (|\theta_G - \theta_B|, |x_G - x_B|, |y_G - y_B|)$. We did not use the norm of the transformation because translations and rotations cannot be compared. This definition was extended to 3D where $T = (\theta_x, \theta_y, \theta_z, x, y, z)$, and $T_e = (|\theta_{x_G} - \theta_{x_B}|, |\theta_{y_G} - \theta_{y_B}|, |\theta_{z_G} - \theta_{z_B}|, |x_G - x_B|, |y_G - y_B|, |z_G - z_B|)$.

This representation of the error in transformation space is used only to analyze how a particular method is exploring the search space but is not intended to evaluate the overall performance of a particular transformation for several reasons: first, we cannot compare two methods that use different transformation spaces (i.e., non-rigid vs. rigid); second, it assigns an accuracy value to both rotations and translation even though they cannot be compared; rotations of points farther away from the origin correspond to larger translations. Finally, the error in computing the fiducial marker registration adds to (i.e., compounds) the errors in the registration methods.

**Partial Hausdorff distance**

The partial Hausdorff distance cannot be used as an indicator of the accuracy between two experiments because different robustness quantiles may have been used, which will bias the results. Despite this problem, we found that there is a correla-

tion between the Hausdorff distance and the other two error measurements described above. In a set of experiments we define the *maximum Hausdorff distance* as the maximum over all images and all repetitions. The *mean Hausdorff distance* is the mean over all images and over all repetitions for a fixed quantile.

## 3.4   Data sets

We used three datasets for three sets of experiments. The first set of experiments was on EPID-EPID registration (2D, unimodal), third on CT-CT (3D, unimodal), and the last one on CT-MR (3D, multimodal). The four datasets are described below.

### 3.4.1   EPID-EPID

Our 2D EPID data set consisted of one reference image and 10 target images with rotations (translations, respectively) up to 20 mm (deg, respectively) from the reference image. These images were generated by changing the position of the phantom on the treatment couch. As mentioned earlier, the EPID images were $512 \times 480$ pixels in size with a pixel dimension of 0.51 mm.

### 3.4.2   CT-CT

We tested PoIRe with a set of 3D Computer Tomography (CT) images. The CT data consisted of approximately 200 slices of 2D images with a slice thickness of 1 mm. Each slice consists of $512 \times 512$ pixels with a pixel dimension of 0.4474 mm.

### 3.4.3 CT-MR

For multimodal 3D registration, we used the datasets provided by the *Retrospective Image Registration Evaluation* (*RIRE*) *Project*. The evaluation website assists in performing an objective, blind evaluation of the accuracy of the registrations performed by PoIRe. We used the first seven patient images of the CT - MR T1 dataset. All the images have a slice thickness of 4 mm. The CT slices have a size of $512 \times 512$ and a pixel size of 0.653595 mm. They have between 27 and 34 slices. The MR slices have $256 \times 256$ pixels with a pixel size of 1.25 mm. The MR images have 26 slices (there were two exceptions with 20 slices). The images were acquired at different institutions as described in [WFW+97].

## 3.5 Experimental setup

We present a comprehensive evaluation of the three algorithmic strategies incorporated in PoIRe. In our study, we set out to compare the performance of the three methods from PoIRe and measure the accuracies achieved. We also evaluated the robustness of PoIRe to noise points. It is important to note that the only preprocessing performed on the point sets was the manual removal of all points corresponding to the markers to avoid biasing the results.

The errors were measured as described in section 3.3.4. To visualize the accuracy of the methods, in our error plots we have drawn a dashed horizontal reference line at 0.51 mm, which is the maximum error that our algorithms were required to achieve. TRE values under this line represent sub-pixel accuracy. This line is not relevant for the rotation component. Each graph shows a 95% confidence interval for the mean of the error. This means that in 95% of the cases, the distribution of the errors is expected to have a mean within the shown interval.

For each of the three algorithmic strategies, there are two parameters that affect the accuracy of PoIRe: the robustness quantile, $q$, and the point sampling rate, $r$. The robustness quantile is associated with the estimated amount of noise in the data sets (See Section 2.3.2). The sampling rate indicates how many points from the original set were randomly selected for each registration experiment. In general, $q$ affects the accuracy, while $s$ affects the efficiency and accuracy of the registration task. Our main objective was to evaluate the impact of the different parameters of PoIRe on its accuracy. Details of the experiments are provided in the next four sections.

The computations in this chapter were done with a 2.8GHz Intel processor with 2GB RAM using Red Hat Linux 3.4.4-2 operating system (1.4GHz Athlon processor with 256 MB RAM for simulated data experiments). The source code was written in C++ and compiled using gcc 3.01. The graphic and web interfaces were written in JAVA 6.

## 3.5.1 EPID-EPID (2D images)

In our first set of experiments, given two EPID images, PoIRe registers these images with subpixel accuracy and with running time averages under 2 seconds. PoIRe is robust enough to work with the low quality images from current commercial EPID technology. While the (deterministic) branch-and-bound was executed only once for each data set, the probabilistic algorithms were repeated 50 times for each one of the 10 target images using the same reference image. This corresponded to 1000 experiments for each choice of parameters. The images had up to 8 fiducial markers, which were used to evaluate the quality of the registration. We used 10 images for our experiments, with one reference image.

**Feature extraction**  In our experiments we used the Canny feature extraction method with a sigma value of 2.5 for the Gaussian mask and a lower threshold of 0.1 for the normalized directional derivatives. The extracted points had as origin the radiation center of the image. The size of the extracted points sets ranged between 1100 and 2000 points.

**Cell size for phase 2 of the Hybrid method**  In this experiment we used an initial cell centered around the approximate answer with a range of $\pm 2$ mm for the translation and $\pm 2$ deg for the rotation. Note that this refers to a region in the transformation space.

**ICP parameters**  We used an initial search radius of 50, a minimum overlap of 70%, a maximum overlap of 100%, a LSQ of 0.0001 and a LSQ limit of 0.00001 and a maximum of 10000 iterations.

**MI parameters**  We used 4 multi-resolution levels. For optimization the Powell method [Pow64]was used for the coarsest resolution and the Levenberg method for the other resolutions. We used the default number of maximum iterations for each resolution, we used 33 bins for the coarsest resolution level and half that number for each finer level.

**Performance evaluation**  The TRE measures were computed for each of the experiments errors using four registration target points (shown in Figure 3.8(a)). We compared our results to the multi-resolution MI approach [TU00] and the *Trimmed ICP* (*TRICP*) [CSK05].

### 3.5.2   CT-CT (3D images)

As explained earlier, the second set of experiments were performed with 3D CT images with 5 fiducial markers and only one image pair.

**Feature extraction**   We used the Deriche algorithm with scale parameter $\alpha = 1$. The point sets had approximately 650,000 points.

**PoIRe parameters**   As explained in Section 3.7.2, the branch-and-bound method was abandoned because its running time (over 12 hours, on our initial tests) for the initial tests made it impractical for clinical purposes. The robustness quantile used ranged from $q = 0.5$ to $q = 0.9$. The sampling probability $s = 0.18$ was used because it was the smallest probability that had a TRE under 1 mm. The CT-CT registrations aimed to show the convergence during the iterations and therefore we did not include a comparison with MI and ICP, which we left for the CT-MR experiments in the following section.

**Performance evaluation**   As before, the TRE measures were computed for each registration using five registration target points.

### 3.5.3   CT-MR (3D images)

The last set of experiments were for intermodal registration and used CT and corresponding MR images from the RIRE data set with one reference image and seven target images.

**Experiment setup**   For MI, we used the results provided by Thevenaz on the RIRE web site. For TRICP and PoIRe we used the following procedure:

1. Use Deriche algorithm to extract edges from the provided training set for both CT and MR T1 images with scale parameter $\alpha = 1$ for both CT and MR [Der93]. The size of the point sets was approximately 200,000 for CT and 70,000 for MR.

2. Run the registration algorithm on the extracted training set with robustness quantiles of 0.5, 0.6, 0.7, 0.8, 0.9 and record the quantile with smallest TRE using the box corners as markers.

3. Select the best quantile and use it for the first seven patients of the RIRE CT-MR data set.

**Performance evaluation** For each experiment, the TRE measures were computed for the hill climbing and the Hybrid methods, for 10 areas of surgical interest in the brain earlier [WFW$^+$97]. The branch-and-bound method was not analyzed because of its large running times. We compared our results to the multi-resolution MI approach by [TU00] and the *Trimmed ICP* (*TRICP*) proposed by [CSK05].

## 3.6 Results

We now present the results of the three sets of experiments mentioned above.

### 3.6.1 EPID-EPID

In our experiments, given two EPID images, PoIRe registers these images with subpixel accuracy and with running time averages under 2 seconds. PoIRe is robust enough to work with the low quality images from current commercial EPID technology.

We compared the performance of the hybrid strategy in PoIRe against MI and TRICP for each of the 10 pairs of images. As seen in Figure 3.9, we can see a peak

| TRE | Branch-and-bound | Hill climbing | Hybrid | TRICP | MI |
|------|------------------|---------------|--------|-------|--------|
| Mean | 0.322 | 0.303 | 0.299 | 0.429 | 2.205 |
| Max | 0.624 | 0.597 | 0.628 | 1.666 | 20.098 |

Table 3.1: TRE summaries for the 2D EPID-EPID image registrations.
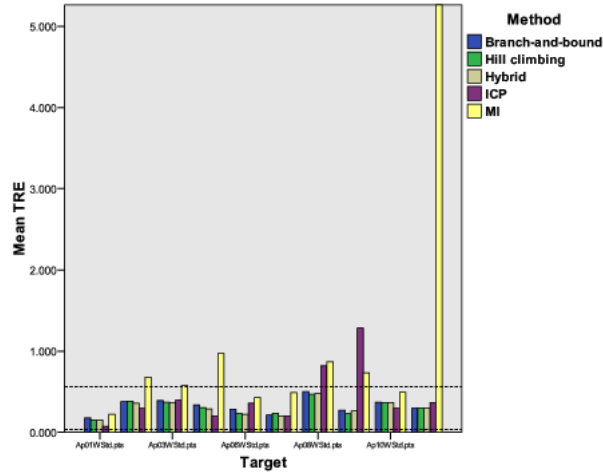


Figure 3.9: Mean TRE comparison for our three methods, TRICP and multi-resolution MI for each pair of images from the 2D EPID data set.

in the TRE for the last image. The graph for the maximum TRE was very similar to the graph for mean TRE (Figure 3.9) and was omitted. Table 3.1 shows the TRE summaries for the 2D registration experiments.

## 3.6.2 CT-CT

Figure 3.10 shows that the algorithm achieves an error under the 1 mm reference line, which is the pixel size in the $z$ direction. Figure 3.11 shows that all the six parameters of the rigid transformation are under the reference line. Additional experiments (data not shown) suggest that the phase-2 of the Hybrid method reduced the TRE and their variance only marginally.
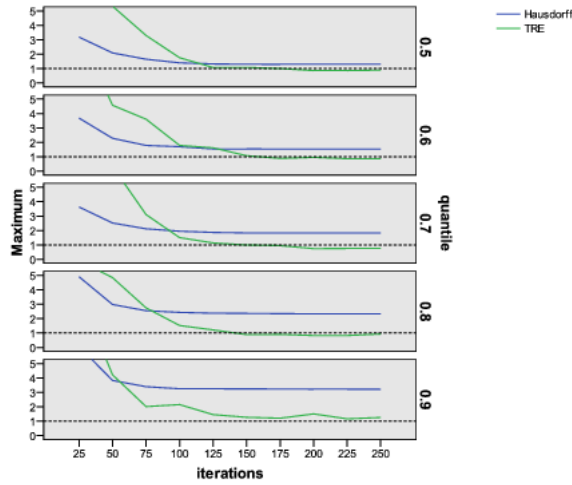
Figure 3.10: Maximum TRE for registration of unimodal (CT-CT) 3D images using the hill climbing method as a function of the robustness quantile. Even with a sampling probability as low as $s = 0.18$, the maximum errors for 10 runs of the algorithm with a quantile value of $q = 0.7$ produces errors under the 1 mm reference line. The strong correlation between the partial Hausdorff distance and the TRE is also clear.
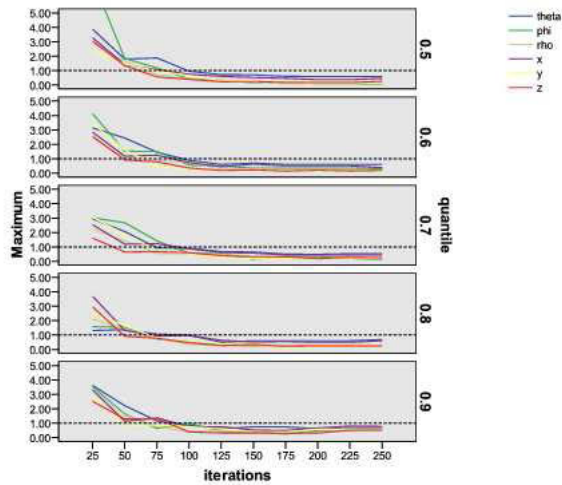


Figure 3.11: Absolute errors in the six parameters of the transformation for registration of 3D images using the hill climbing method. This figure correspond to the same setting as in Figure 3.10 but shows the errors in transformational space. Again, a quantile $q = 0.7$ has the smallest errors.
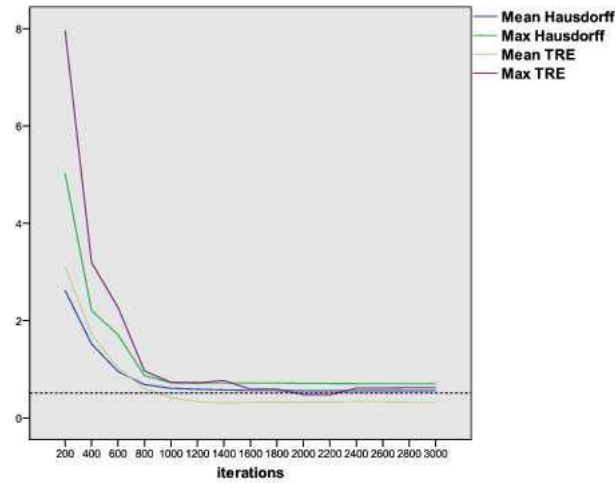
Figure 3.12: Error analysis for registration of 2D images using the branch-and-bound method with a robustness quantile $q = 0.8$. Errors below the horizontal line correspond to sub-pixel accuracy.
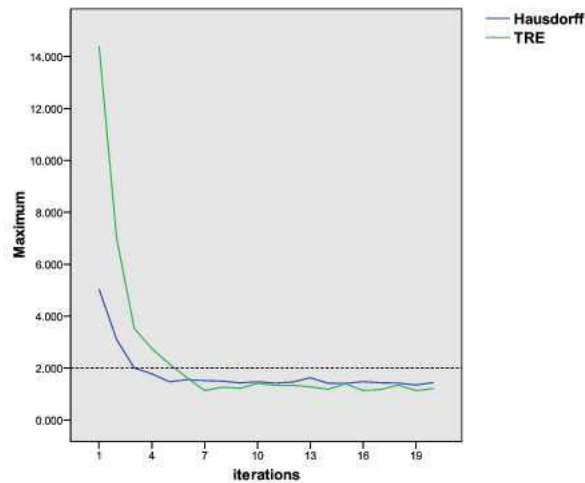


Figure 3.13: Maximum error analysis for registration of 2D images using the stochastic hill climbing method (as phase 1 of the Hybrid method) randomly sampling 20% of the points. Since this is only phase 1 of the hybrid method, the goal was only to achieve reasonable accuracy (2 mm) and not sub-pixel accuracy. At this phase we are only interested in the maximum error, because we want to ensure that even the worst case reached such reasonable accuracy. Fifty repeats were performed.
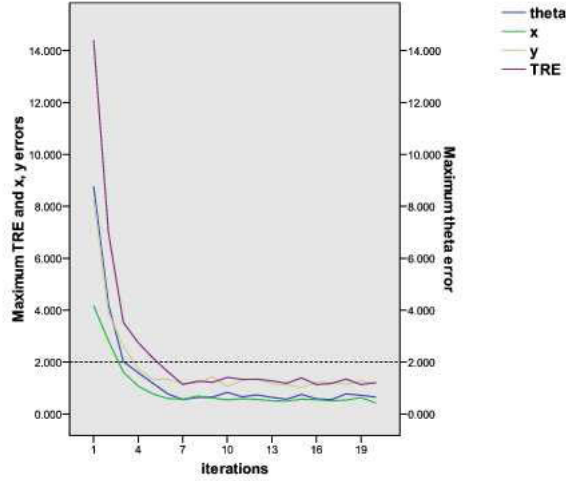
Figure 3.14: Error analysis (Translational, rotational and TRE) for registration of 2D images using the level 1 of the Hybrid method (stochastic hill climbing method) randomly sampling 20% of the points. The x, y and theta errors represent the error vector as described in Section 3.3.4. As in Figure 3.13, we are only interested in achieving reasonable accuracy and use the maximum to guarantee that even the worst registrations achieve such accuracy. Fifty repeats were performed.
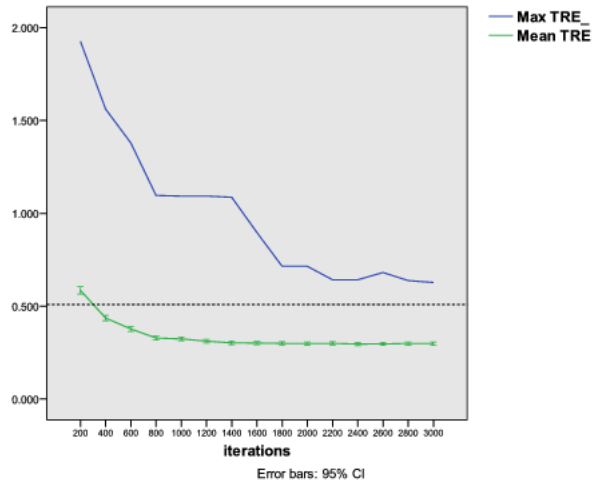


Figure 3.15: Error analysis for registration of 2D images using the branch-and-bound (as phase 2 of the Hybrid method). No sampling and no repetitions performed.

69

(a) 0 iterations

(b) 100 iterations

(c) 600 iterations

(d) 1200 iterations
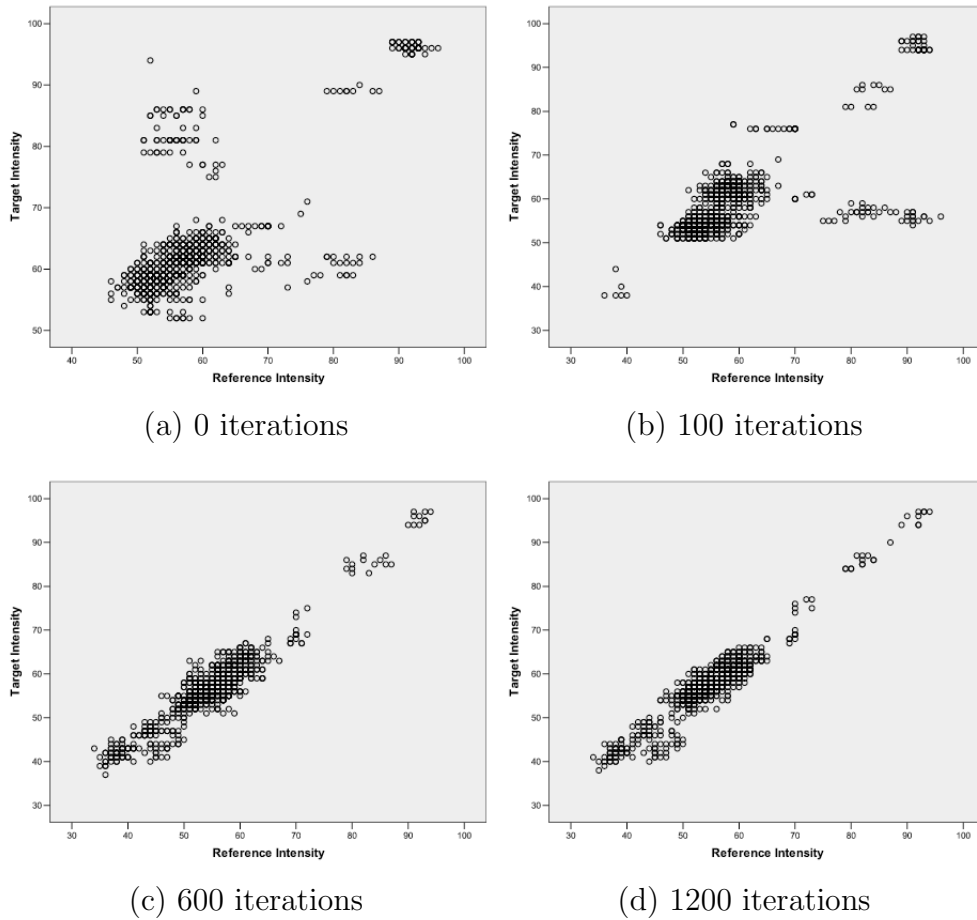
Figure 3.16: Reference and target intensity values at different iterations of the Branch-and-bound method. The X axis represent the intensity values of extracted feature points in the reference image. The Y axis represents the intensities of the matched feature points in the target image. For each reference point, its nearest neighbor at different stages of the registration is computed. Even though PoIRe makes no use of pixel intensities it is interesting to see the increase of correlation between reference and target intensities as out algorithm evolves. At iteration 1200 the registration error (in transformation space) respect to the marker-based ground truth was $(0.05, 0.09, 0.14)$

| TRE | MI | TRICP | PoIRe |
|--------|-------|--------|-------|
| Mean | 1.363 | 5.613 | 2.758 |
| Max | 2.654 | 20.405 | 8.529 |
| Median | 1.266 | 4.633 | 2.100 |

Table 3.2: TRE summaries for the CT-MR T1 (RIRE) image registrations.
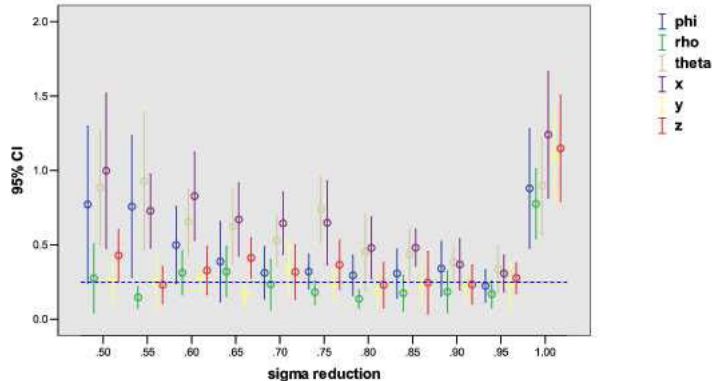
Figure 3.17: Error analysis for registration of 3D images (using the hill climbing method) as a function of the variance reduction parameter. The error decreases consistently all the way to 0.98 (very slow reduction of variance), increasing for a value of 1, which means no reduction at all.

### 3.6.3 CT-MR

Table 3.2 shows that our mean and maximum errors are smaller than the corresponding TRICP errors. We are under 1 mm over the MI accuracy (given the ground truth error). In all experiments, the errors using PoIRe are under 4 mm, which is the separation between slices in both the CT and MR images.

## 3.7 Discussion

In this Section we discuss the results for the 2D, the 3D unimodal and the 3D multi-modal registrations

### 3.7.1 The 2D Case

For EPID-EPID registration we highlight the following topics:

**Branch-and-bound method**   The partial Hausdorff distance (Eq. (2.1)) has one important parameter: the robustness quantile $q$. There is no reason to expect that

the optimal choice of $q$ is the same for different data sets. This parameter is related to the amount of noise in the point sets that needs to be ignored. Therefore it is dependent on the image modality and the feature extraction method used, and perhaps also on the subject. Choosing large values of $q$ assumes that the amount of noise is small and the risk of using noise points for the registration increases. Choosing small values of $q$ assumes that there is a large amount of noise and the risk of discarding good points that provide significant tissue information increases, with the consequent loss of valuable information for accurate registration. We tested quantile values from 0.5 to 1 with increments of 0.05. The best accuracy for the 2D data set using the branch-and-bound method was achieved by using a quantile value of $q = 0.8$. TRE increased slightly for both $q = 0.75$ and $q = 0.85$ suggesting that 20% noise points may be intrinsic to their data sets. This assumption held for all ten target images for all three methods, giving the smallest TRE for $q = 0.8$. It is important to note that for $q = 0.7$ and $q = 0.75$ (data not shown), the algorithm did find solutions with acceptable errors within a small number of iterations, which were however abandoned as the algorithm progressed.

The branch-and-bound algorithm finds a transformation with subpixel accuracy under 1000 iterations, taking roughly 5 seconds having a TRE very close to 0.51 mm, as shown in Figure 3.12.

**Hill climbing method** Figure 3.13 shows the maximum registration errors for the stochastic hill climbing algorithm with a sampling probability of 0.2 for both point sets and a quantile $q = 0.5$. The hill climbing algorithm converges on the eighth iteration. The time required for all the twenty iterations is under 0.5 sec. This makes stochastic hill climbing a promising candidate for fast registrations in 3D.

It is important to understand how PoIRe performs with point sets of different sizes. Furthermore, it is useful to know if all the feature points identified by the feature extractor are important. So we considered the option of sampling from the feature point sets. We investigated the effect of sampling on the accuracy of PoIRe, its running times, and the optimal choice of robustness quantiles.

Initially we tried the same optimal quantile from Figure 3.12 (where the optimal robustness quantile was 0.8 and no sampling was used) but had large TRE errors. Experimentally, we found that the value of the optimal quantile gets smaller for smaller sampling probabilities. Intuitively, one would expect that the optimal robustness quantile does not change with uniform sampling because the expected ratio of good versus noise points is also likely to be preserved in the sample. However, sampling in the two images is performed independently and therefore the likelihood that both points from a matched pair are sampled is lowered. Hence the need to lower the robustness quantile for smaller sampling rates. In Figure 3.13 it can be seen that for a sampling probability of 0.2 , the optimal quantile for hill climbing is 0.5.

The confidence interval for the mean error is small and similar in all the images of the data set. This suggests that once the best quantile is found for a modality and for an experimental setup, it will work for other images with the same setup.

We have shown that for 2D EPID images, TRE of under 0.322 mm are achieved. Given the nature of the EPID images and the high degree of noise points extracted from EPID images, it is clear that the method is robust to noise.

**Hybrid method**    Figure 3.13 shows the maximum TRE and Hausdorff error using stochastic hill climbing method as the first phase of the proposed two-phase Hybrid method. Even though we sampled only 20% of the points, the maximum TRE goes

from 14 mm to under two mm very quickly. The robustness quantile used was $q = 0.5$. We found experimentally that with small randomly sampled point sets, the number of good matching points gets reduced and therefore the choice of quantile has to be reduced as well. We are using this estimate to reduce the search space for the branch-and-bound. We also reduce the search to a box of size 2 (mm, deg) as indicated by the reference line. However, to make sure that we are bounding the search space properly, we need to look at the errors in the transformation space.

Figure 3.14 shows the errors in the transformation space of the hill-climbing algorithm (used as phase 1 of the Hybrid method) when randomly sampling 20% of the points. It shows the correlation between the TRE and the parameters of the 2D rigid transformation. In this case, we can see that coincidentally, the TRE is close to the translation error in the $y$ coordinate. The x, y and theta errors represent the error vector as described in Section 3.3.4. We can see that all the maximum errors for the registration parameters are well under the 2 mm/deg reference line and we can use these results as good starting points for the bounded branch-and-bound search.

**Branch-and-bound as phase 2 of the Hybrid method**  Figure 3.15 shows the error analysis (TRE) for registration of 2D images using branch-and-bound as phase 2 of Hybrid method. The performance is fast but we were right over the reference line but the confidence interval was close to half pixel; the branch-and-bound tightens the 95%-confidence interval. Sub-pixel accuracy is achieved before 400 iterations making it very quick although the maximum error is larger than that of the pure branch-and-bound.

**Correlation between Hausdorff distance and TRE**  Figure 3.12 shows the mean and maximum TRE and Hausdorff error for the branch-and-bound algorithm.

The Pearson correlation between TRE and Hausdorff (for all 10 images and 15 iteration measurements) was 0.915 with a significance level of 0.01. The final mean TRE for the branch-and-bound was 0.322 mm and the maximum TRE was 0.50, being marginally under the 0.51mm (1 pixel) reference line. Having this correlation is an interesting feature of the Hausdorff distance. Note that since the algorithm optimizes the partial Hausdorff distance, it always decreases monotonically.

**Correlation between the intensities of matched pixels**   Even though our method makes no direct use of intensity values, we can see in Figure 3.16 the correlation of intensity values of the pixels corresponding to pairs of matched points between the reference and target images at different stages of the branch-and-bound. There is a clear increase in this correlation as the algorithm progresses. At iteration 1200 the registration error (in transformation space) with respect to the marker-based ground truth was $(0.05, 0.09, 0.14)$.

**Mutual information Error**   As seen in Figure 3.9, we can see a peak in the TRE for the last image. The corresponding EPID image and its extracted point set are shown in Figure 3.8(b). For this particular image, the mean and maximum FRE were 0.284 and 0.457 respectively so it cannot be argued that the large MI error is due to out-of-plane rotations. Even if we ignore this troublesome image, the mean TRE would be 0.60 and the max 1.11, both well over the mean error for the Hybrid method (including the 0.262 FRE uncertainty from the ground truth). This unstable behavior makes MI undesirable for a system that uses EPID images. Errors with TRICP are under the interval for the mean but also have a large maximum TRE of 1.666 mm.

**Fiducial error**   Using SVD, the fiducial marker registration was computed with a mean FRE of 0.26 mm with a standard deviation of 0.08 and a maximum value of 0.48 mm. We used 5 markers that were common to all images and then discarded the one that consistently had a large residual error, leaving the ground truth to be determined by 4 markers.

### 3.7.2   The 3D Case unimodal

For CT-CT registration we highlight the following three topics:

**Effect of sampling**   While sampling does lower the quality of the registration, the goal was to determine the levels of sampling that would maintain the errors at acceptable levels.  As mentioned before, the Hausdorff distance is sensitive to the absence of a one-to-one correspondence between feature points. In other words, because of heavy sampling, the probability of sampling both points of a pair of "matched" points is small, leaving many unmatched points, which the algorithm then needs to discard as noise points. This is evident in the 3D CT-CT image registration experiments. Even though the mean error values have sub-pixel accuracy, the standard deviations for translations are above 0.1 mm.  This correspondence problem was not present in EPID-EPID registrations where standard deviations for translations were approximately 0.01 mm because no sampling was used.

Our first approach was to sample very small point sets and have them as input for the branch-and-bound and hill climbing components of PoIRe. Sampling 10% of the points and using the branch-and-bound algorithm had an error of 0.5 degrees for rotations and 0.3 mm for translations. It takes over 360,000 iterations and a run time of 2.2 hours. This makes branch-and-bound unsuitable for the clinical applications. For the hill climbing, the maximum error for all runs was under 0.5

degrees and 0.4 mm. It required 500 iterations with a CPU time of 6.6 minutes. Given the difference in running times for the two algorithms, the branch-and-bound algorithm was not considered for further analysis.

Figure 3.10 shows the TRE error, along with the Hausdorff distance, for hill climbing registration of 3D images for five different robustness quantile values. A sampling probabilities of 0.18 was used. We used very small samples trying to achieve a quick estimate of a good start point for the hybrid approach and were interested in defining an appropriate initial bound for the search space of the branch-and-bound method that would encompass *all* the results from the hill climbing method (not on the average, but even in the worst cases). Therefore, the plots show the *maximum* errors for 10 runs of the algorithm for each image. Figure 3.11 shows the error is the transformation space. For $q = 0.7$, the errors were under the reference accuracy line. With a sampling rate of 0.07 the CPU time was 3.3 min. With 0.18 it increased to 12 min.

**Hybrid approach**    The results from section 3.7.2 using the hill-climbing algorithm were used as starting points for the branch-and-bound with an initial cell of size 1 mm (one pixel).

Unfortunately, phase 2 of the Hybrid method showed only marginal improvements over the starting points (data not shown).

**Experimental parameter settings for hill climbing**    The parameters of the stochastic hill climbing were fixed for all the experiments presented here. A population of size 100 with an offspring percentage of 20% were set from the beginning. We found that a Hebbian learning coefficient of 0.9 produced the smallest errors. Figure 3.17 shows the registration errors (in transformation space) as a function of the variance reduction factor which reached a minimum at 0.98.

Because of the blind nature of the RIRE project, we could not capture the accuracy of the registrations *during* the execution of the algorithms.

**Fiducial error**  Using SVD the fiducial marker registration was computed with a mean FRE of 0.26 mm and a maximum value of 0.42 mm. We used the 5 fiducial markers to compute the ground truth.

### 3.7.3   The 3D Case multi-modal

For the CT-MR registration of the RIRE data sets we discuss the following topics:

**PoIRe compared to ICP**  We found that PoIRe performed much better than ICP for this experiment. The mean TRE was five times smaller and the maximum TRE was ten times smaller than the mean using ICP. This supports our claim that PoIRe is more robust to noise and less sensitive to starting conditions. Slight distortions in the MR data seem to affect the accuracy of ICP. This large difference in accuracy suggests that PoIRe's hybrid method together with the Hausdorff distance provide a robust framework.

**PoIRe compared to MI**  Mutual information has been a standard for MR-CT registration. Our mean TRE was twice as large than the mean TRE using mutual information. This is not surprising according to our preliminary results that showed PoIRe not suitable for multi-modal registration. A better approach for this type of registration might be the approach suggested in Chapter 4.

**Fiducial error**  The fiducial error for the RIRE data is 0.55 mm for CT-MR [WFW⁺97].

## 3.8 Conclusions

An efficient and robust method for performing multimodality image registration was presented. It involved novel features such as the weighted Hausdorff distance and the stochastic hill climbing as a local optimization technique in the branch-and-bound algorithm. In addition, the performance was improved by introducing distance based priority, and multi-resolution feature extraction.

This algorithm was applied to the practical problem of patient positioning in radiotherapy. Our experiments demonstrated subpixel accuracy for intramodality image registration (i.e. EPID-EPID) and pixel accuracy for multimodality image registration (i.e. DRR-EPID). Mean computation times for both cases were under 2.5 seconds using an Athalon 1.4 GHZ processor running Linux 7.3 operating system, and offers a potential clinical application for automated patient positioning in radiotherapy.

The work presented here evaluates the accuracy of a feature-based medical image registration method. We present results for registration of 2D EPID images and 3D CT scans of an anatomical phantom. The results show that branch-and-bound methods produced registration with sub-pixel errors in 2D and sub-pixel mean errors in 3D. Stochastic hill climbing produced even better results both in 2D and 3D and the hybrid method produced low errors with running times under 5 sec in 2D and 5 min in the 3D case. Sampling turned out to be a good option for faster registrations with slightly larger error variances but we found that different sampling probabilities required the estimation of appropriate values of the robustness quantile. We also clearly demonstrated the need for robustness. The non-robust version with a quantile of 1 (meaning 100% of the points are inliers) performed poorly. With a quantile of 0.8 the errors were considerably reduced achieving the subpixel accuracy

with errors under the 0.25 mm threshold line. We showed that subpixel accuracy can be achieved by PoIRe in a radiotherapy setting with running times under 2 seconds.

The branch-and-bound was unable to take advantage of starting from a transformation close to the global optimum because, in general, the lower bounds were very small and very few cells were killed based on the similarity of the starting point.

We found that both point-based methods (PoIRe and TRICP) performed better than MI for the EPID-EPID registrations. MI was better for the 3D case.

# CHAPTER 4

## IMAGE REGISTRATION USING GABOR FEATURES

In this chapter we turn our attention to the improvement in accuracy of the point-based rigid registration using features related to local texture captured by the complex circular Gabor filter. Circular Gabor filters have also been reported in literature as a successful tool for image classification. In this particular application, we apply the Gabor filters on multi-modal (DRR-EPID) images and apply them to the image registration with applications to patient positioning in cranial radiotherapy.

The weighted partial Hausdorff distance can be readily modified to include additional information about the images. Our first attempt to modify the Hausdorff distance was to include the gradient information from the Canny edge detector into the computation of the weights. Our initial intuition was that noise points would have higher gradients than bone edges. Our idea was to penalize those noise points. These initial experiments were performed for 2D EPID-EPID registration. Because of the nature of EPID images, the bone contours all turned out to have a very similar gradient making unsuccessful any attempt to use them as a discriminating function. It was not until we extracted information from a *neighborhood* around each point of interest that we were able to have a successful weighting function for 2D images [PP07].

We saw that our Hausdorff approach outlined in Chapter 3 produced large errors when registering multimodal images. Using only spatial information led to point correspondences that had no anatomical correspondence due to the different responses to the edge detectors. In this chapter we use additional information from the image pixels to guide the computation of a Hausdorff distance, with the hope that the nearest neighbors according to the modified distance measure would have better anatomical correspondence. For this analysis we use a set of training image

pairs to characterize the correlation between different features. The ground truth in these experiments is computed using marker-based registration. We apply the ground truth to the target point sets in the training set and select the mapped points that lie within a neighborhood $\epsilon$ of a point in the reference point set.

## 4.1   Introduction

Several variants of the weighted Hausdorff distance function have been proposed to incorporate additional information into the definition of point similarity. Gesu and Starovoitov proposed a weighted Hausdorff function that includes the gray scale difference between two images $I$ and $F$ to achieve reliable registrations [GS99]. It is defined as follows:

$$d_E(\mathbf{x}, \mathbf{y}; I, F) = \sqrt{\omega_1((x_1 - y_1)^2 + (x_2 - y_2)^2) + \omega_2(I(\mathbf{x}) - F(\mathbf{y}))^2}, \qquad (4.1)$$

where $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ are location vectors of pixels in the two images $I$ and $F$, respectively, with intensities $I(x)$ and $F(y)$, respectively; $\omega_1$ and $\omega_2$ are arbitrary constants.

Instead of comparing image intensities directly, Yang et al. used normalized gradient differences to provide illumination invariance by reducing the influence of very dark or very bright regions in the images [YLC04]. The function used by them to compare pixel coordinates and intensities is:

$$E(\mathbf{x}, \mathbf{y}; I, F) = \left| \frac{\nabla I(\mathbf{x})}{\max_{(m,n) \in W_x} |\nabla I(m,n)| + c} \cdot \frac{\nabla F(\mathbf{y})}{\max_{(m,n) \in W_y} |\nabla F(m,n)| + c} \right|,$$

where $\nabla I$ and $\nabla F$ are the gradient vectors of images $I$ and $F$ respectively, $W_x$ and $W_y$ are windows centered at the points $\mathbf{x}$ and $\mathbf{y}$ respectively, and $c$ is a small constant.

Incorporating the function $E$ in the distance function (Equation 4.1) gives the following function [YLC04]:

$$d'_E(\mathbf{x}, \mathbf{y}; I, F) = \sqrt{\omega_1((x_1 - y_1)^2 + (x_2 - y_2)^2) + \omega_2(1 - E(\mathbf{x}, \mathbf{y}; I, F))^2}. \quad (4.2)$$

Including this information into the directed partial Hausdorff distance (Equation 2.1) results in the gradient-weighted Hausdorff function

$$H_k(A, B; I, F) = \max_{a \in A} \min_{b \in B} d'_E(x, y; I, F), \quad (4.3)$$

where $A$ and $B$ are two given point sets.

Similarly, the gradient-weighted symmetric modified Hausdorff is defined as:

$$S(A, B; I, F) = \max \left\{ \frac{1}{|X|} \sum_{x \in X}^{H} d_{Y,F}(x : I), \frac{1}{|Y|} \sum_{y \in Y}^{H} d_{X,I}(y : F) \right\}, \quad (4.4)$$

where

$$d_{Y,F}(x : I) = \min_{y \in Y} \left\{ d'_E(x, y; I, F) \right\}$$

and

$$d_{X,I}(y : F) = \min_{x \in X} \left\{ d'_E(x, y; I, F) \right\}.$$

Yang et al. set $\omega_1$ and $\omega_2$ to a value of one, assigning equal importance to the spatial domain as well as the intensity domain and showed that including the gradient information improved the recognition rates for face recognition applications. Equation (4.1) is a weighted distance function that uses the Euclidean distance. In addition to the Euclidean distance, the authors also provided weighted versions for the city-block and chessboard distances [YLC07].

We attempted (separately) using pixel intensities and gradient information in the weighted Hausdorff distance for EPID-DRR registration. Both experiments failed because of the multi-modal nature of the registration and the non-existence of a linear relationship between intensities or gradients. We decided to evaluate other

local descriptors and used the Gabor filters to extracted relevant information from the images that could be incorporated in the weighted Hausdorff distance.

### 4.1.1 Gabor Features

Gabor functions became known in computer vision because of their ability to model the receptive field of simple cells in the visual cortex. These cells fire up after receiving an input of parallel lines with a specified orientation and frequency. Although this fact was reported previously, there has been no further development of this property in the literature. Gabor filters are widely used in pattern recognition and machine vision applications because of their spatial localization and selectivity in terms of orientation and frequency; another relevant property of these filters is their invariance to changes in illumination [KKK06].

The traditional Gabor function is defined as:

$$G(x, y) = g(x, y) \exp(2\pi i(f_x x + f_y y)), \qquad (4.5)$$

where $(f_x, f_y)$ is the center frequency, and the Gaussian modulating function $g$ is defined as:

$$g(x, y) = \frac{1}{2\pi\sigma^2} \exp(-(x^2 + y^2)/2\sigma^2). \qquad (4.6)$$

The traditional Gabor function is not rotation invariant, which is a problem for image registration. Gabor features have been used for image classification [AGP06]. Each image is associated with the mean and variance of the norm of the Gabor response for a set of scales and orientations. To provide invariance for rotation, they shifted the feature vectors circularly starting with the largest mean, which is called *dominant direction*. This is based on the assumption that rotating textures in such a way that the dominant directions match is necessary for comparing similarity between textures. They showed an improvement in the classification accuracy

compared to Wavelet Statistical and Wavelet Co-occurrence features. It was proved earlier that image rotation in space is equivalent to circular shift of the feature vector [ZWIL00].

Zhu et al. used the circular Gabor function [ZTL04]. They also introduced an odd Gabor filter for edge detection. They presented a weighted robust version of the modified Hausdorff distance (MHD) (see [DJ94]) by adding a robustness quantile (see [HR93]). The weights are based on a binomial function of the distance of a feature point to the center of the image and aim to decrease the influence of the background. A set of $l$ Gabor features and one spatial feature (normalized distance to the center of the model, robust to scaling) are used in the computation of the Hausdorff distance. They compute the resulting Hausdorff distance for every pixel in the image using as coarse location the pixel where it is minimized. At this location, a window W is used find a finer solution based on the spatial distance of the correspondence between points defined by the Hausdorff distance. Robustness to Gaussian noise, occlusion, scale and rotation are demonstrated experimentally.

Xu et al. used a Circular Gabor Filter (CGF) [SC06] to overcome the problem of orientation selectivity in feature extraction as an alternative to Traditional Gabor Filters (TGF) to extract features for K-means clustering for segmentation of biological data. A bank design method is presented also following the one presented by Jain et al. [JF91]. The CGF has only two parameters, while the TFG has four. The authors emphasize that creating a comprehensive jet of CGF responses takes fewer responses than TGF, making CGF computationally more efficient. It was shown experimentally that the frequency band of simple cells in the visual cortex is approximately 1 octave [Pollen 83]. Given that the frequency bandwidth from frequency $f_1$ and $f_2$ is $\log(f_2/f_1)$ we set the frequencies to be of the form $2^n C$. The bank design method involves creating a bank with frequencies one octave apart as

$\sqrt{2^S}/N$, where $N$ is the image width in pixels, and $S = 0, 1, ..N/4$. The $\sigma$ is derived from the equation:

$$B = \log_2 \frac{\pi F \gamma \sigma + \sqrt{\ln 2/2}}{\pi F \gamma \sigma - \sqrt{\ln 2/2}}$$

and corresponds to $3\sqrt{\ln 2/2}/\pi F$. Here the computation gets simplified in the CGF because we are assuming $\gamma = 1$.

Wang and Chua used a bandwidth $B = 0.55$ and claimed it to be an appropriate choice for face recognition [WC05b]. It is still an open problem to determine the appropriate bandwidth for multi-modal image registration. Wang and Chua used 3D Gabor filters with $F = 2^{i+2/2}\pi$, $i = 2, 3, ..., 6$, and used the Hausdorff distance on the vector of responses. To compare Gabor features they built a sphere around the point of interest and computed the histogram of Gabor responses using 4 bins between the maximum and minimum values of both the real and imaginary part of the responses and computed the squared sum of the differences between the histogram components.

In a later work, a jet of Gabor responses was used by Wang and Chua for face recognition, invariant to facial expression and perspective [WC06b].

Wang and Tang presented a method that used Gabor features and Bayesian learning for image classification [WT03]. It builds a *jet* or vector of responses for 5 scales and 8 orientations of the Gabor filter. Fiducial points are extracted as described by Wiskott et al. [WFKvdM97]. Each fiducial point is annotated with this Gabor jet. Wang and Tang's contribution is in the use of Bayesian classification. Wiskott et al. introduced graph *bunches* which are responses at the same node with different expression/illumination [WFKvdM97].

Liu et al. extracted Gabor features at select facial locations by dividing the facial image into key and assisting points based on masks generated by PCA and using coarse subsampling for assistant points and a finer sampling for key points

to improve the recognition rate [LLS04]. Key points are selected based on the 25% largest magnitudes for Gabor responses.

We selected circular Gabor filters mainly for their invariance to rotations [ZTM02, SC06]. These filters are defined as the product of a complex exponential sinusoidal and a Gaussian modulating function $g$ [KKK02].

The Gabor filters are defined in terms of a complex exponential, with specific frequency and orientation, and a modulating Gaussian function. Their responses are similar to the receptive field profiles observed in cortical simple cells, whose localization, orientation, and frequency-selectivity properties [JF91] have been exploited, among others, in face recognition tasks [WC05b]. In a further development, Zhang et al. [ZTM02] propose a rotation invariant circular Gabor filter for classification purposes, defined in 2D as:

$$G(x, y) = g(x, y) \exp(2\pi i f \sqrt{x^2 + y^2}), \tag{4.7}$$

where $f$ is the central frequency of the circular Gabor filter, and the Gaussian modulating function $g$ is defined as in Equation (4.6).

The Fourier transform of the CGF which looks like a ring is:

$$F(u, v) = \frac{\sqrt{2\pi}}{2} \alpha \exp(-(\sqrt{x^2 + y^2} - F)^2 / 2\alpha^2).$$

In 3D, the spherical Gabor filter is defined as:

$$G(x, y, z) = g(x, y, z) \exp(2\pi i f \sqrt{x^2 + y^2 + z^2}) \tag{4.8}$$

and the modulating function is defined as:

$$g(x, y, z) = \frac{1}{(2\pi)^{\frac{3}{2}} \sigma^3} \exp(-(x^2 + y^2 + z^2) / 2\sigma^2) \tag{4.9}$$

The Hausdorff distance between two vectors of complex responses $A$ and $B$ was defined by Zhu et al. [ZTL04] as follows:

$$H(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|.$$
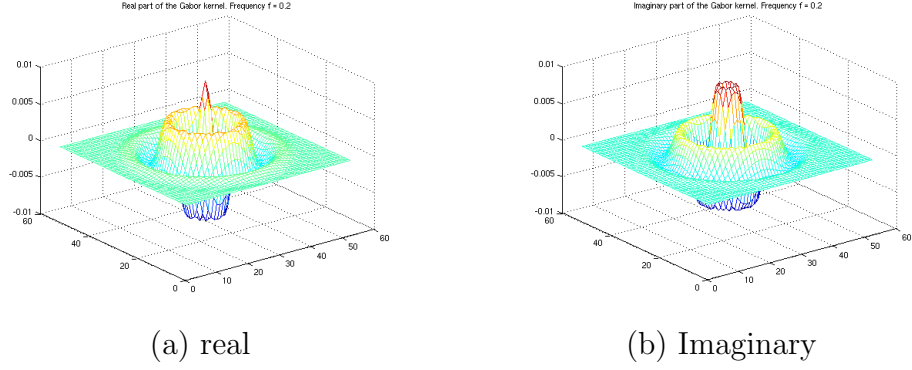
87

(a) real
(b) Imaginary
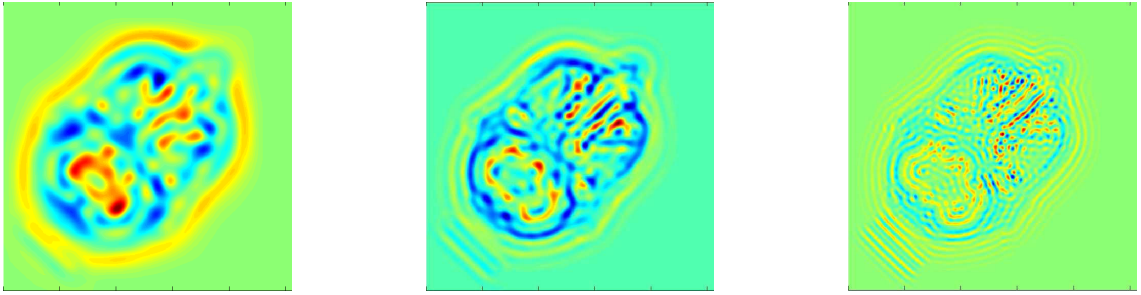
Figure 4.1: Gabor filter with frequency 0.2



Figure 4.2: Responses for different frequencies

We define the weighted Hausdorff distance as:

$$H_w(A, B) = k_{a \in A} \min_{b \in B} dist_w(a, b), \qquad (4.10)$$

where

$$dist_w(a, b) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + w \sum_{i=1}^{n} (x_{g_i} - y_{g_i})^2} \qquad (4.11)$$

and $n$ is the number of Gabor features.

## 4.1.2 Canonical correlation analysis

Canonical correlation analysis is a method used to weight factors in a statistical analysis predicting pairs of variables with high correlation. This method allows us to compare two groups of variables by computing pairs of orthogonal factors which

link the two variable groups and, as in PCA, produces an ordering of the correlations between the eigenvalues.

Let $Z$ be a matrix of centered measurements (with mean 0) divided into two groups. We want to find vectors $a$ and $b$ that allows us to maximize the correlation between $Z_1$ and $Z_2$.

$$Z = [Z_1 Z_2],$$

The covariance matrix $V$ of Z is a symmetric matrix that we can divide in 4 groups

$$V = \frac{1}{n} [Z_1 Z_2]^\top [Z_1 Z_2] = \begin{pmatrix} V_{11} & C_{12} \\ C_{21} & V_{22} \end{pmatrix},$$

where $V_{11}$ and $V_{22}$ are the variance-covariance matrix of the elements of $Z_1$ and $Z_2$, respectively, and $C_{12}$ is the correlation between elements of $Z_1$ and $Z_2$ with $C_{21}^\top = C_{12}$.

We want to find pairs of factors $\{u_p = Z_1 a_p, v_p = Z_2 b_p\}$ that are uncorrelated within their own groups but that maximize the correlation $\text{corr}(u_p, v_p)$ referred to as their *canonical correlation*, and given by

$$corr\,(u, v) = \frac{a^\top Z_1^\top Z_2 b}{\sqrt{a^\top Z_1^\top Z_1 a}\sqrt{b^\top Z_2^\top Z_2 b}}.$$

If we normalize $a$ and $b$ by the metric given by the variance covariance matrices, $V_{11} = Z_1^\top Z_1$ and $V_{22} = Z_2^\top Z_2$, respectively, we have $a^\top V_{11} a = b^\top V_{22} b = 1$ and we can write the correlation as follows:

$$corr(u, v) = a^\top C_{12} b.$$

With the substitution $x = \sqrt{V_{11}} a$ and $y = \sqrt{V_{22}} b$ we can express the canonical correlation as:

$$corr(u, v) = x \left[ \sqrt{V_{11}} C_{12} \sqrt{V_{22}} \right] y,$$

If $G = \sqrt{V_{11}} C_{12} \sqrt{V_{22}}$, then

$$X^\top G Y = \Sigma,$$

that is to say the singular value decomposition is given by:

$$G = X \Sigma Y^\top,$$

where the canonical correlations, which are the only non zero elements of $\Sigma$, are the singular values of the rectangular matrix $G$.

If we maximize the canonical correlation function using Lagrange multipliers we can see that the Lagrange parameters correspond to the correlation. We can see that maximization of the correlation

$$\phi = a^\top C_{12} b - \frac{1}{2} \mu \left( a^\top V_{11} a - 1 \right) - \frac{1}{2} \mu' \left( b^\top V_{22} b - 1 \right)$$

is achieved by setting the partial derivatives to zero as follows:

$$\frac{\partial \phi}{\partial a} = 0 \quad \Leftrightarrow \quad C_{12} b - \mu V_{11} a = 0, \text{ and}$$

$$\frac{\partial \phi}{\partial b} = 0 \quad \Leftrightarrow \quad C_{21} a - \mu' V_{22} b = 0$$

Then, we have $C_{12} b = \mu V_{11} a$ and $C_{21} a = \mu' V_{22} b$.

Premultiplying the first equation with $C_{21} V_{11}^{-1}$ and substituting in the second we have

$$V_{22}^{-1} C_{21} V_{11}^{-1} C_{12} b = \lambda b, \text{ and}$$

$$V_{11}^{-1} C_{12} V_{22}^{-1} C_{21} a = \lambda a.$$

To find the relationship between the Lagrange parameters and the canonical correlation we premultiply the partial derivative equations with $a^\top$ and $b^\top$ respectively giving us:

$$a^\top C_{12} b - \mu a^\top V_{11} a = 0 \quad \Leftrightarrow \quad \mu = a^\top C_{12} b = corr(u, v), \text{ and}$$

$$b^\top C_{21} a - \mu' b^\top V_{22} b = 0 \quad \Leftrightarrow \quad \mu' = b^\top C_{21} a = corr(u, v)$$

## 4.2   Method

This section describes how we applied the Gabor filters for feature extraction and how we reduced the dimensionality of the Gabor response vector using canonical correlation analysis.

We first consider pixel how the features behave when two images are registered using the given ground truth. We start with pixel intensities as the feature of interest. We see in Figure 4.3(a) that the scatter plot of the intensities of a DRR and the corresponding registered EPID images is fairly wide with points congregated roughly along the diagonal and with a correlation coefficient of 0.883. This is because pixel intensities vary widely for different modalities. We extracted Gabor features from the images. We filtered each image with a Gabor kernel and computed the response at each point for 10 different frequencies of the Gabor filter. The correlations of Gabor features for mapped points were in general very low and only a few variables showed significant correlation. In Figure 4.3(b) we see the plot of features A4 and B4 with the largest correlation of 0.913 and in Figure 4.3(c) we see the plot of features A5 and B5 with the second largest correlation of 0.905.

To solve this problem we propose the use of canonical correlation analysis which produces a set of orthogonal vectors for each image and represents the Gabor features as a linear function in that orthogonal system. The new pairs of variables correspond to eigenvectors and their correlation corresponds to the eigenvalues of the system. The eigenvalues can be sorted (i.e., decreasing correlation) and only the significant ones can be selected, thus reducing the data representation to a small number of variables.

One of the problems of Gabor features is that the Gabor function at different frequencies does not form a (orthogonal) basis to represent the input data.
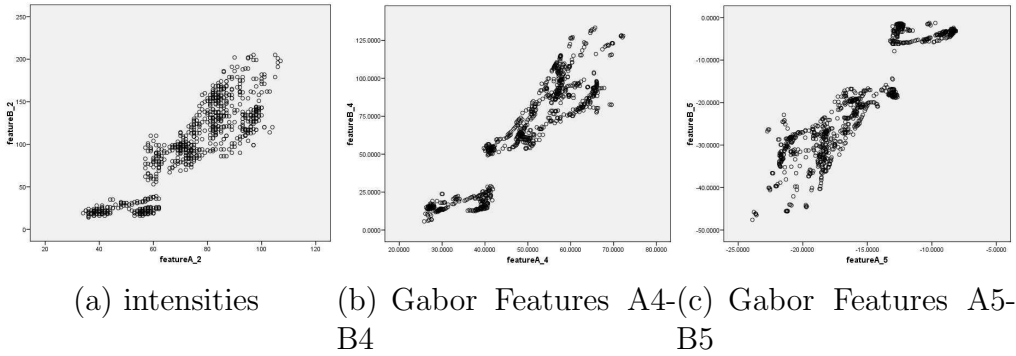
(a) intensities  (b) Gabor Features A4-B4  (c) Gabor Features A5-B5

Figure 4.3: Feature extraction



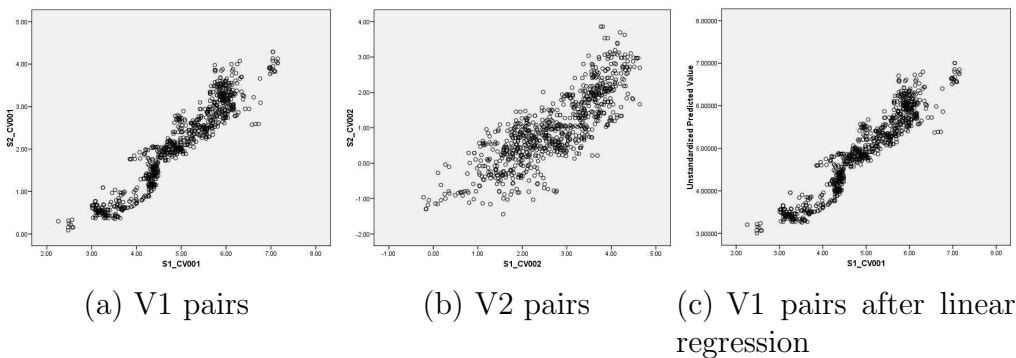(a) V1 pairs  (b) V2 pairs  (c) V1 pairs after linear regression

Figure 4.4: Canonical correlation analysis pairs

For the new set of variables, we see in Figure 4.4(a) that the largest correlated pair of variables with the largest correlation has a correlation of 0.952 and that the pair with the second largest correlation of 0.822 (see Figure 4.4(b)). These variables are independent which allows for consistent comparisons when associated with spatial information for registration. Although the variables in Figure 4.4(a) have the largest correlation, we can see that they are not scaled, which is desirable in the computation of the Hausdorff distance. To scale the data statistically we use linear regression to produce the data in Figure 4.4(c).

| $w_a$ | $w_b$ |
|-----|-----|
| NRM | NRM |
| NRM | INV |
| INV | NRM |
| INV | INV |

Table 4.2: Mapping linear function for gradient to compute weights

## 4.2.1 Gradient as local descriptor

Our first systematic set of experiments was to try to help points with lower gradient to win the nearest neighbor search in the Hausdorff distance and for the high gradient points to get pushed down in the selection of the $K^{th}$ largest of such distances. The Hausdorff distance is redefined as follows:

$$H(A, B) = \max_{a \in A}^{k} w_a \min_{b \in B} w_b \cdot dist(a, b), \tag{4.12}$$

where $w_a$ and $w_b$ are the *normalized* (NRM) gradient values for the points in set $A$ and $B$. This means that the weights will have a value between 0 and 1. Because we wanted high gradients to have low ranks, for $w_b$ we *inverted* (INV) their values, meaning that values of 0 would correspond to higher gradients and 1 to lower ones. Our experiments did not produce any good results. We then proceeded to check all the combinations of choices for the two weights as shown in Table 4.2. We also considered ignoring the gradient information and assigning a weight of 1 to all the points in the set. We called this the *unweighted* (UNW) case.
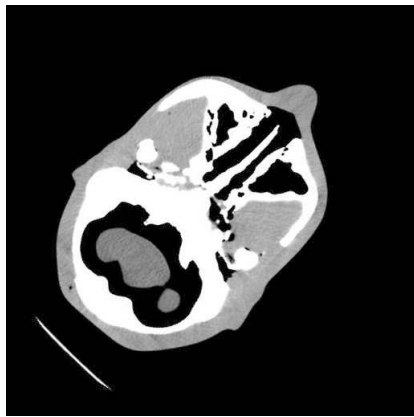
## 4.2.2 Experimental setup

For this particular study we used kernels of size 27, based on a 10-frequency model, leading to the computation of a vector of 10 complex responses for each extracted point. The resulting registration algorithm had a computation time of 2.70 seconds,
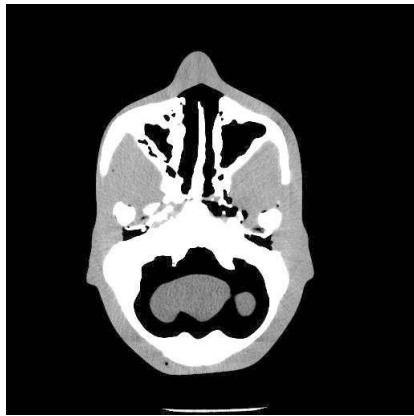
(a) Reference image
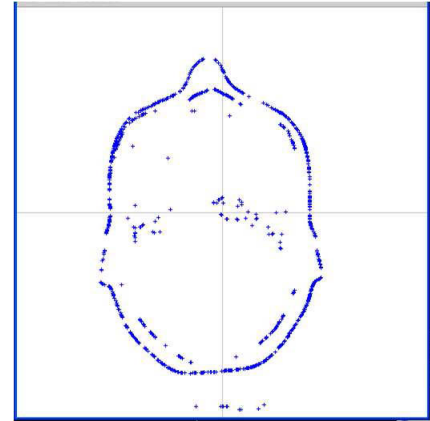


(b) Reference point set



(c) Target image
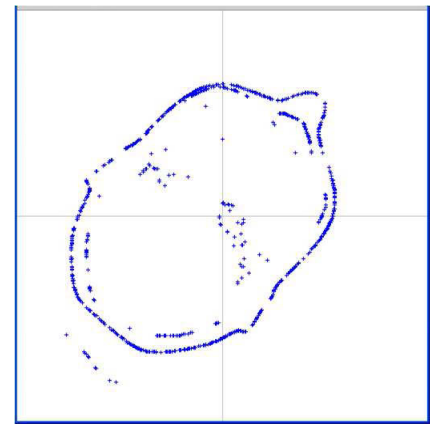


(d) Target point set

Figure 4.5: Initial experiment

from which 2.10 seconds were used for computing the Gabor responses, while the remaining 0.6 seconds were invested in registering the point sets. These registration times (averaging 2.7 seconds) are based on a feature extractor that selected roughly 1400 feature points selection, using a 1.4 GHz processor with 256 MB of main memory.
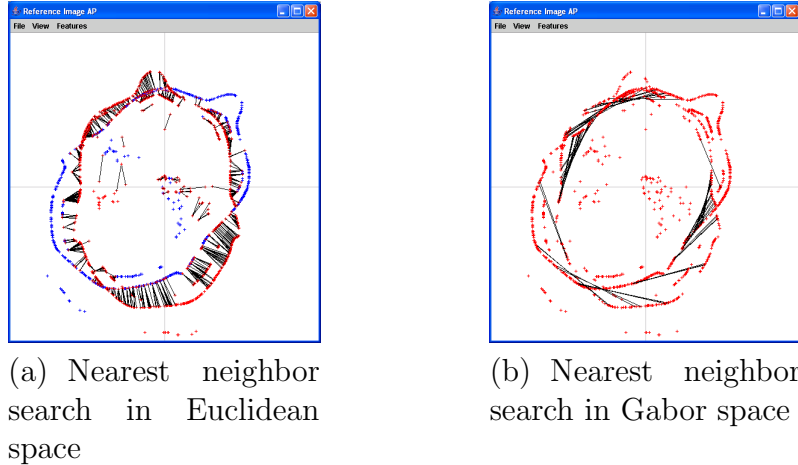
(a) Nearest neighbor search in Euclidean space

(b) Nearest neighbor search in Gabor space

Figure 4.6: Nearest neighbor search

## 4.2.3 Point-based registration

We use the partial Hausdorff distance that is computed in the space of responses to the circular Gabor filter. Thus, each point of interest in the reference and target images generates a circular Gabor filter, which computes a vector of complex responses to different Gabor kernels. Each kernel corresponds to using different frequencies and scales of a rotation invariant Gabor function, which are then used to determine corresponding regions in the images to be registered. Responses to a circular Gabor filters is reported in the literature as a robust method for image classification. The partial Hausdorff distance measure considered in this study to compute the distance between images was defined in Chapter 3 as:

$$H(A, B) = \max_{a \in A} \min_{b \in B} dist(a, b) \tag{4.13}$$

where $k$ $(0 < k \leq |A|)$ is the robustness parameter that corresponds to the number of outliers rejected during the computation of this distance. This variable can be defined in terms of the quantile $q$, $0 < q \leq 1$, by the relation $k = q \cdot |A|$.

The Hausdorff distance is computed in the spatial domain, obtaining for each point in image A the corresponding nearest neighbor in image B; often, this process

95

does not match the same anatomical features. The approach used here departs from the above approach in the following aspects: i) we perform the Hausdorff distance computation in the Gabor space to determine a set of corresponding pairs of points from each image that minimizes such a distance; ii) from the nearest neighbor search in Gabor space, we obtain point pairs that either correspond to the same anatomical structure, or that show similar textures; iii) we exploit the rotation invariance of the circular Gabor filters to define corresponding points, even in situations where the images are rotated.

### 4.2.4 Gabor features for registration

As described in Section 4.2.1, the first set of experiments failed to produce good results. We then decided to take a closer look at the behavior of the weights with the ground truth. Note that, in our experiments, the ground truth was provided by the SVD-based registration of markers extracted from the image. We used the Pearson correlation coefficient on the final pairs that were within a distance $d$ (initially set to half a pixel width). We set out to find the Gabor responses (or a combination of the responses) that had a high correlation for pairs of points that ought to be paired under the "ground truth" registration. Then we would have to see if the resulting weighted Hausdorff distance could discriminate points and select good pairs and reduce the registration errors. We used gradients and intensity values and found out that their correlation for the nearest neighbors after registration does have a very low correlation coefficient. As shown in Figure 4.3, high correlation was found only with the first variable of the canonical correlation analysis performed on the set of Gabor responses.

## 4.3 Results

Experiments were performed for EPID-DRR registration with one reference DRR and 12 EPID images. EPID-DRR registration is very difficult because of the resolution quality of the EPID images. We tried several different weights including the norm of the complex vector, the phase/angle representation. However the only one that showed reasonable performance was the canonical correlation analysis on the set of real and imaginary response pairs for each frequency.

We computed ten Gabor responses for DRR and EPID images and added this responses to each point in the extracted point sets. The edge detection parameters were relaxed to include a large set of edge points. Each response has a real and a imaginary component, resulting in a vector of 20 features for each point. Canonical correlation enforces that the inter-category correlations to be high while the intra-category correlations remain low. The problem with having high intra-category correlation is that the model will attempt to compensate creating opposing values to simulate one uncorrelated variable, resulting in a poor analysis. Because of the image characteristics, responses with high inter-category correlations were not possible unless large kernel sizes were used. The kernel used was of size 53x53, extracting information from a large number of neighboring pixels. It turned out that the imaginary response of the lower frequencies produced larger inter-category correlations. Intra-category correlations were still large but we proceeded to use the method and used only the first variable from the canonical correlation analysis, which had a correlation larger than 0.9. We performed the registration using an additional coordinate for the variance and weighted it appropriately. Figure 4.8 shows the registration errors as a function of the weight. For a weight of 3.25 the mean errors were under the 0.51 pixel accuracy line. For the unweighted case (weight =

0) the errors in the $y$ coordinate were close to 0.7 mm, which is unacceptable.

Our test images were obtained from a cranial phantom, and included a reference EPID image, ten additional EPID images with rotations and translations, and one DRR obtained from the 3D CT projection that aligns the fiducial markers with those observed in the EPID reference.

The Gabor-based method described above is quite robust in registering high energy scatter EPID images, which are characterized by elevated levels of noise. In terms of pixel accuracy, we obtained a 95% confidence interval under 0.51mm, which represents a substantial improvement over previous feature-based registration experiments [PP07].

Regarding the weighted Hausdorff distance weight parameter $w$ (see Eq. 4.11), we obtain robustness (a stable solution) for $w$ values close to 3.25. As shown in Fig. 4.7, registration errors increase for either lower or greater $w$ values. We also obtain an individual Gabor feature correlation of 0.918 (compared to 0.839 in the case of intensities). Canonical correlation analysis offered a first variable correlation of 0.943 for the EPID-DRR case.

## 4.4 Experiments

These experiments aim to analyze the effects of several feature extraction parameters and help setting up the 3D experiments to conclude the work on the Gabor features-based effort.
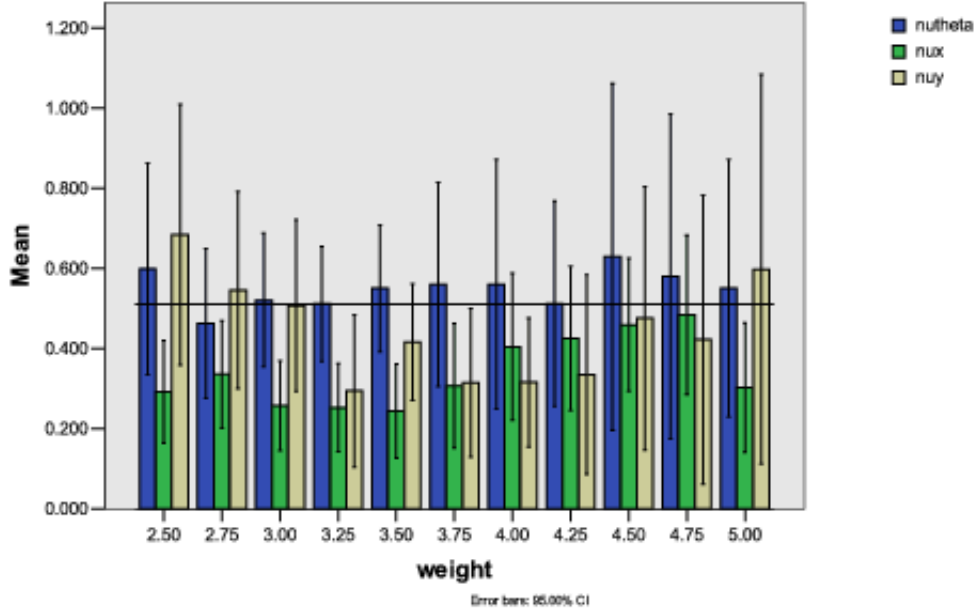
Figure 4.7: Registration errors for rotation, translation in $x$, and translation in $y$, as a function of the weight parameter in equation 4.11

## 4.4.1 Experimental parameters

**Feature extraction parameters**

A fixed set of points was used for reference and 10 fixed sets were used as targets. Each point in these sets was annotated with information about the texture in its neighborhood. We compute the complex Gabor feature responses for 10 frequencies $(0.1, 0.2, 0.3, \cdots, 1)$.

We used two parameters to study these responses:

1. Kernel size: Refers to the number of pixels used in each coordinate to compute the Gabor responses. We have computed the Gabor responses using kernel sizes of 11, 21, 31, 41, 51, 53, and 61 pixels. Given that both DRR and EPID images have the same spatial resolution, we used the same kernel size for reference and targets.
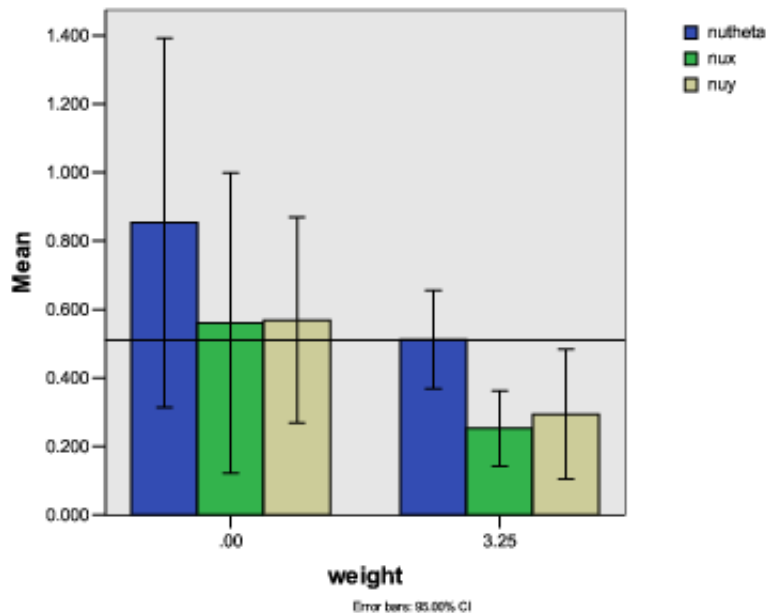
Figure 4.8: Comparison of registration errors for rotation, translation in $x$, and translation in $y$. The bars at $w = 0$ correspond to the errors for an unweighted distance computation [PP07], whereas the averages at $w = 3.25$ illustrate the best error reduction obtained from our EPID-DRR registration.

2. Frequency scale: Since the frequency range of 0.1 to 1 was not guaranteed to be the best to extract texture information, the vector was scaled using values of 0.25, 0,5, 1, 1.5 and 2. The responses of DRR and EPID are independent and therefore all possible combinations of frequencies and sizes needed to be tested.

**Applying canonical correlation analysis**

The correlations between the 10 extracted variables coming from the CT and EPID images are quite low. We used canonical correlation analysis to create a set of highly correlated variables. This analysis requires the points in the images to be paired. We mapped the target to the reference using the ground truth and selected the nearest neighbor to each target point to select its match.

1. We considered points that were within 0.25 (more restrictive), 0.5, 0.75 and 1 mm (more relaxed) from their matched point as "good matches".

**Registration parameters**

1. Robustness Quantile: We used quantile values of 0.5, 0.6 and 0.7.

2. Weight: Weight values from 0 to 4 with increments of 0.5 were used. The same weights were used for all the annotated response coordinates.

3. Number of variables from the canonical correlation analysis added to each spatial point: Currently only the first one was used, but it would make sense to try to use up to the fifth largest.

**Evaluation parameters**

The TRE measure evaluation was performed using 4 target points. The mean TRE error for a registration is the mean TRE for these 4 points.

**Conclusion**

We have 7 kernel sizes, 4 frequency scales for DRR, 4 frequency scales for EPID, 4 NN distances, 3 quantiles, 9 weights and 5 canonical correlation variables to study. This is equivalent to more that 60000 experiments, each one performing 10 registrations.

## 4.4.2 Round I: Narrowing down the best parameters

**Experimental parameters**

For this set of experiments, the frequency values used were limited to 0.25 and 0.5 (4 combinations for the 2 images), the NN distance was fixed to 0.5 mm, and only the first correlated variable was used. Also was excluded the kernel size of 51 because of

its closeness to 53. (The value of 53 was kept to verify and reproduce the previous experiment.)

**Experimental results**

For all experiments, the quantile values of 0.6 and 0.7 had large TRE values. Figure 4.9 shows the mean and maximum TRE. The mean TRE error for the unweighted case (weight = 0) was close to 2 mm. We selected 8 cases for which the mean TRE error was close to the 1 mm reference line (Marked with a circle in Figure 4.9). The smallest TRE corresponded to a kernel size of 31 pixels, a frequency scale of 0.5 for both reference (DRR) and target (EPID), and a quantile of 0.5. These parameters were used for the next experiment.

Although only the best parameters were used in the next experiment, it is important to notice that the quality of the registration is similar for values close to the best solutions. This suggests that the weighted distance is robust and the quality of the registration will not be affected by small changes in the weight value.

### 4.4.3   Round II: Using the best parameters

**Experimental parameters**

Using the parameters that resulted in the smallest TRE from Round I above, $q$ was set to 0.45 and 0.5, and the NN distance was set at 0.5 mm to determine the "good" point matches, with values 0.25, 0.5 (computed previously), 0.75 and 1.

**Experimental results**

Figure 4.10 shows the two experiments that resulted in the lowest TRE achieved. The first one was performed with a NN distance threshold of 0.25, a weight of 3.5
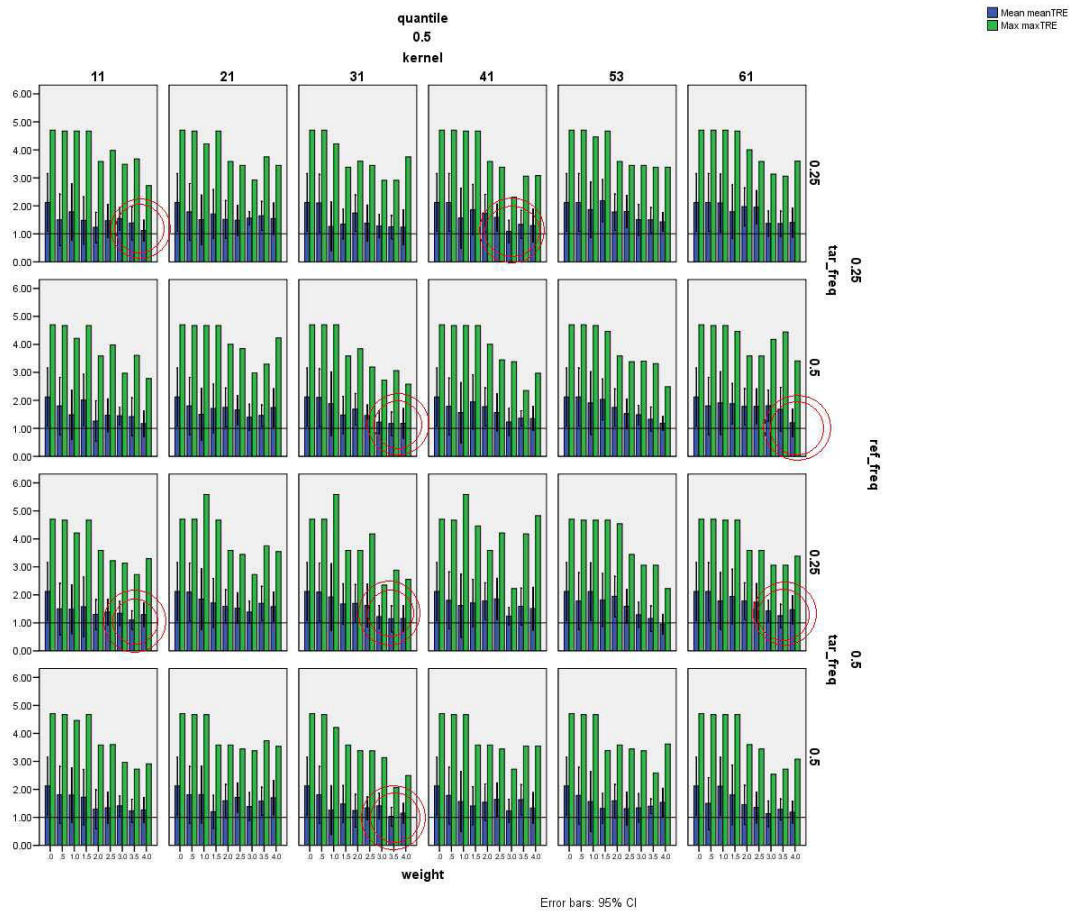
Figure 4.9: TRE as a function of kernel size and filter frequency. Quantile values is 0.5. Other quantile values (0.6 and 0.7) had large TRE for all cases.
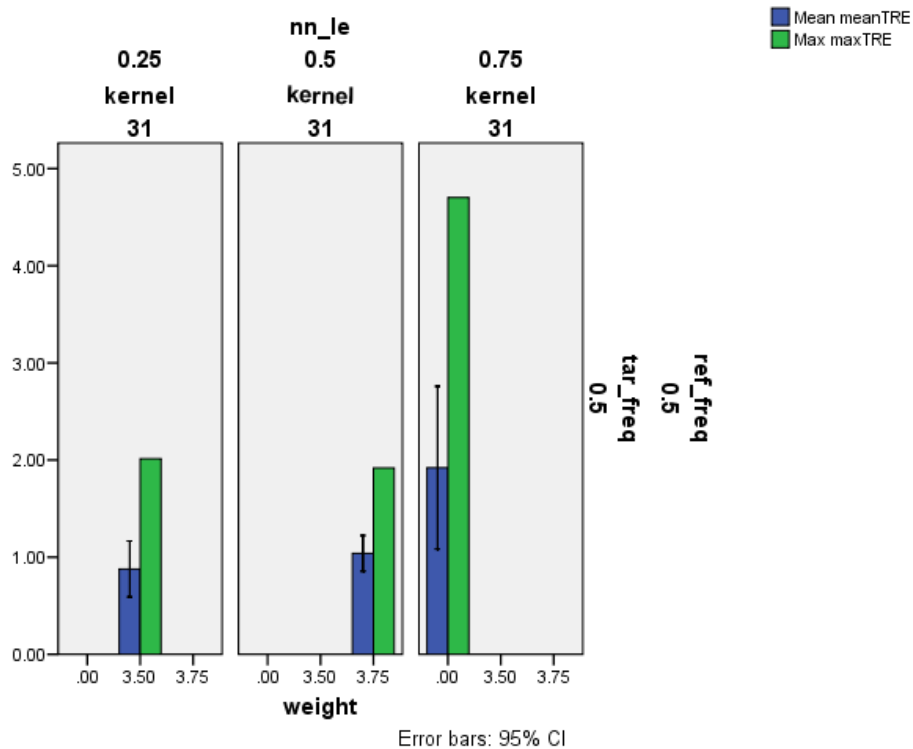
Figure 4.10: Best (two) TRE values for the second experiment. The third panel shows the unweighted case. The TRE using mutual information has a mean value of 2.66 and a maximum TRE of 5.34.

and a quantile of 0.45. The second one was performed with a NN distance threshold of 0.5, a weight of 3.75 and a quantile of 0.5. The third panel shows the unweighted case with a mean TRE close to 2 mm and and maximum TRE close to 4.5 mm. Both weighted cases had a maximum TRE close to 2 mm and the 95% confidence interval for the mean under 1.2 mm with a mean under 1 mm for the first one.

## 4.5   Conclusions

We showed in this chapter that the weighted Gabor registrations reduced considerably the maximum and mean TRE for EPID-DRR registration. We showed that our approach reduces the mean and maximum TRE errors for EPID-DRR registration

by more than 50%.

The proposed feature-based method offers an accurate and robust method for the registration of cranial radiotherapy images, exploiting anatomical features derived from EPID and DRR images. Experimental results suggest that our procedure is appropriate for patient position verification, and is therefore a valid alternative to invasive surgical implant of fiducial markers.

The use of circular Gabor filters coupled with the canonical correlation analysis allows to reduce the error in DRR-EPID registration, and these results suggest that the method can be extended for registration of 3D CT and 2D EPID images.

Additional experiments are also required to evaluate the number of variables from the canonical correlation analysis. The registration might benefit from adding additional (orthogonal) variables in the model at a cost of performance.

The good behavior of the Gabor features for registration suggests that additional experiment using other local descriptors in a similar fashion will allow to improve the quality of the registrations. Our study also indicates that this procedure can be generalized to other imaging modalities.

CHAPTER 5

## NON-RIGID POINT-BASED SHAPE REGISTRATION

The two previous chapters focused on global, rigid transformations, which are useful in medical imaging for images of the brain, pelvis and large bones. For other regions of the body or for intra-subject registration, the rigid model might not be able to produce good registrations. Non-rigid local deformations are common in medical imaging, particularly in applications such as in respiratory-gated PET imaging. Since the chest expands and contracts during respiration, a rigid transformation cannot adequately register images taken at different stages of the respiratory cycle. This chapter presents a new point-based non-rigid registration method that uses the Hausdorff distance and the free form deformation (FFD) model. Instead of focusing on general images, we will focus on shapes, which are simplified digital images where each pixel value is either 0 or 1.

## 5.1   Introduction

Several algorithms have been proposed to solve the non-rigid shape registration problem. Local non-rigid models include optical flow [AWS00], Thin Plate Splines (TPS) [Boo89] and Free Form Deformations (FFD) [SP86]. The search strategy usually corresponds to a variational approach where a gradient descent-like method is used. Point-based methods usually require a determination of the point correspondence between the point sets.

Non-rigid registration transformations can be represented by an optical flow deformation field. The main drawback of optical flow is that it does not necessarily preserve the topology and coherence of shape after deformation. The Thin Plate Spline (TPS) model is widely used but its convergence is very sensitive to the established point correspondences. TPS interpolates these correspondences instead

of approximating them, therefore the method is not robust to mismatches. Free Form Deformations (FFD) are very popular as well, and involve embedding an object in a space and deforming it by deforming the space. One of the most common interpolating basis function is the cubic B-spline. We use this basis in our FFD model.

In the previous chapters we used the Hausdorff distance measure, implemented using the Kd-Tree approach [Ben90]. In the Kd-Tree approach, each computation of the Hausdorff distance involves computing the nearest neighbor in $B$ for every point in $A$. This defines a one-to-many correspondence relationship between the points in $A$ and the points in $B$ defined by the nearest neighbor relation. In this sense it resembles what is used by the ICP strategy [BM92]. The Distance Transform (DT) approach on the other hand, provides every point in space with the distance to its closest point in $B$ without explicitly providing the information of which point in $B$ is actually the nearest neighbor. Consequently, the Hausdorff distance is computed without explicitly determining correspondences between the point sets. For point-based non-rigid registration, this is a novel approach. Methods that explicitly determine correspondences between points use local descriptors and the accuracy of the registration relies on their ability to discriminate matching pairs accurately.

Implicit representations have been successfully used for shape registration on variational frameworks [HPM06, PRR03]. These representations are generic in the sense that they handle data in different dimensions and are robust to shape perturbations and noise. We only make use of free form deformation, which, at its coarsest level, corresponds to a projective transformation. Our method does not require the computation of correspondences between points and is thus less likely to get trapped in local minima.

This chapter presents a novel hierarchical method for global and local regis-

tration. We use a shape representation (see [Bor88]) where the reference shape is represented as a distance function (implicit representation) and the target shape is represented as a point set. We use a hierarchical Free Form Deformations (FFD) model based on B-splines and recover the optimal FFD parameters by minimizing a function of the local Hausdorff distance using stochastic hill climbing. At its coarsest level, the FFD model only accounts for a global perspective transformation and allows non-rigid mappings at finer levels. We achieve registration results similar to those achieved with intensity-based methods without requiring expensive evaluation of pixel similarities. The use of a robust Hausdorff distance prevents misregistration of noise points and the free form registration provides $C_2$ continuity. The point-based evaluation avoids the use of a variational approach, resulting in faster registrations.

We show experimentally the accuracy of our model and also outline its limitations. We will see that contour-based registration by itself does not provide a solid ground for non-rigid registration but embedded in a hierarchical approach combined with a robust point similarity measure provides a good approximation of the registration.

## 5.2  Method

In order to perform point-based non-rigid registration we deform the target point set using a Free Form Deformation (FFD) model that produces a deformed position for each point in the target point set by deforming the space that contains both the reference and target point sets. A modified version of the Hausdorff distance is minimized to find the optimal space deformation parameters. To compute this distance efficiently, we used the Distance Transform representation for the reference

and the node-list representation for the target point sets [Bor88]. The result is a robust non-rigid method that registers point sets using a smooth space deformation. See Figure 2.1 for the distance transform representation of a point set.

## 5.2.1 Free Form Deformation (FFD)

Let the rectangular domain $\Omega = \{(x, y) \, | l_x < x < h_x, l_y < y < h_y\}$ be the space such that $A \cup B \subset \Omega$. The lattice of control points $P = \{P_{m,n}\} = \{P^x_{m,n}, P^y_{m,n}\}$, $m = 1, \ldots, M$, $n = 1, \ldots, N$, overlaid on the domain $\Omega$ determines a mapping for every point in $\Omega$ including the points of $A$ and $B$.

For any point $\omega = (x, y) \in \Omega$, the mapping $T(\omega)$ is a function of the control points and is defined as

$$T(\omega) = \sum_{k=0}^{3} \sum_{l=0}^{3} B_k(u) B_l(v) P_{i+k,j+l}, \tag{5.1}$$

where $1 \le i < M$ and $1 \le j < N$, are the number of lattice cells to the left (below) the cell containing $\omega$ along the $x$ and $y$ direction respectively, and $u$ and $v$ represent the normalized value ($0 \le u, v \le 1$) of $x$ and $y$, respectively, relative to the grid that contains the point $\omega$. They are defined as $i = \left\lfloor \frac{x-l_x}{h_x-l_x} (M-1) \right\rfloor$, $j = \left\lfloor \frac{y-l_y}{h_y-l_y} (N-1) \right\rfloor$, $u = \frac{x-l_x}{h_x-l_x} (M-1) - \left\lfloor \frac{x-l_x}{h_x-l_x} (M-1) \right\rfloor$ and $v = \frac{y-l_y}{h_y-l_y} (N-1) - \left\lfloor \frac{y-l_y}{h_y-l_y} (N-1) \right\rfloor$. Each $B_k(u)$ function corresponds to the $k^{th}$ cubic basis function of the B-spline defined as follows:

$$
\begin{aligned}
B_0(u) &= (1-u)^3 / 6, & B_1(u) &= (3u^3 - 6u^2 + 4) / 6, \\
B_2(u) &= (-3u^3 + 3u^2 + 3u + 1) / 6, & B_3(u) &= u^3 / 6.
\end{aligned}
\tag{5.2}
$$

The FFD provides a one-to-one mapping and a deformation field that is smooth and $C_2$ continuous because it is defined as the sum of $C_2$ continuous basis functions.

## 5.2.2 Local Hausdorff Distance

In chapter 3, we have used the Hausdorff distance successfully for rigid point-based registration. Using this distance directly in our non-rigid framework leads to degenerate results as shown in Figure 5.1(a). To produce valid results we have found that the computation of the Hausdorff distance in a neighborhood around each control point $P_{m,n}$ leads to stable solutions as shown in Figure 5.1(b).

As seen in Figure 5.2, a given control point $P_{m,n}$ has an effect on a limited area of the space around it. We denote this subset as $G(P_{m,n})$ and will be referred to as the *support* of $P_{m,n}$. Given a point set $A$, only points in $A \cap G(P_{m,n})$ get affected by changes in the control point $P_{m,n}$. Points closer to the control point are affected more acutely as per the $T$ function (see Equation 5.1). Therefore in order to evaluate a given control point, we want to penalize mismatching points close to it more than mismatches in points far away that receive almost no influence from this control point. Similarly, we want to encourage movements that keep the points close to the starting position. This motivates the following novel modified Hausdorff distance presented below:

$$H_{LOCAL}(A, B, P_{m,n}) = K^{th}_{a \in A \cap G(P_{m,n})} \frac{\beta \Phi_B^2(a)}{B_D\left(w\left(\sqrt{d(a, P_{m,n}) \cdot d(T(a), P_{m,n})}\right)\right)}, \quad (5.3)$$

where $\Phi_B(a)$ is the distance from point $a$ to its nearest neighbor in $B$ and $B_D(d)$ is defined as

$$B_D(d) = \begin{cases} B_1(d), & d \leq 1 \\ B_0(d), & d < 1 \leq 2, \\ 0, & \text{otherwise,} \end{cases}$$

where the cubic B-spline basis $B_0$ and $B_1$ are defined in Equation (5.2), $T$ is the point map function defined in Equation (5.1), and is defined as $w(x) = x/|\sigma|$, where $\sigma = \left(\frac{h_x - l_x}{M-1}, \frac{h_y - l_y}{N-1}\right)$, and $\beta$ is a scaling parameter.

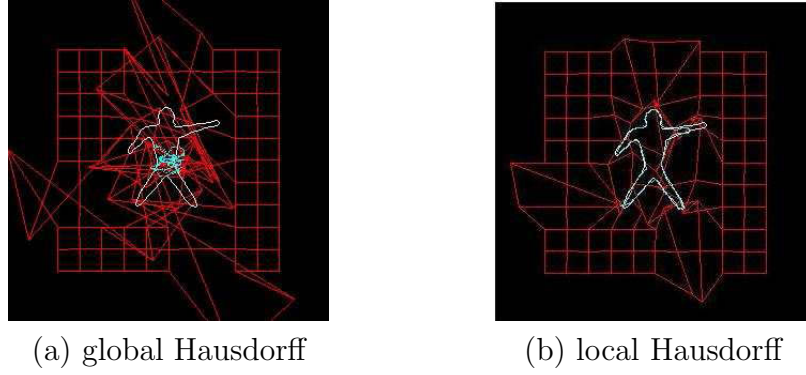(a) global Hausdorff                    (b) local Hausdorff

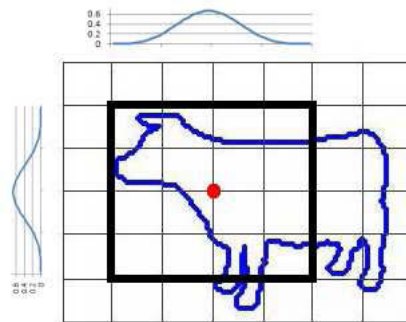Figure 5.1: Global and local Hausdorff distance



Figure 5.2: The support $G(P_{m,n})$ (shown as a black box with thick border) for control point $P_{m,n}$ (shown in red)

### 5.2.3 Hierarchical FFD

We use a multi-resolution method that uses hierarchical refinement based on the notion that each one of the FFD basis functions (Equation( 5.2)) can themselves be described as a linear combination of smaller versions of the same basis functions [FB88]. At the coarsest level, a lattice with $M = N = 2$ accounts for projective registrations and, therefore provides a global registration as outlined in [HPM06, PRR03]. On the other hand, at finer levels, it is expected for some control points $P_{m,n}$ to have an empty set $A \cap G(P_{m,n})$ allowing the algorithm to ignore such control points and focus on the control points for which this set is not empty. Also, the implicit assumption that displacements are local around the $G(P_{m,n})$ area will prevent the finding of good deformations that involve a wider region. Figure 5.3 shows the hierarchical FFD used for shape registration.

### 5.2.4 Optimization

The function to minimize is given by the following:

$$H_{OPTIM}(P, A, B) = \alpha(H(A,B))^2 + (1 - \alpha)\sum_{n=1}^{N}\sum_{m=1}^{M} H_{LOCAL}(P, A, B_{n,m}),$$

where $\alpha$ is a parameter that controls the impact of the global Hausdorff distance $H(A, B)$ over the sum of the local Hausdorff distances.

We used stochastic hill climbing (Section 2.5.4) to explore the search space looking for the control points $P_{m,n}$ that minimize $H_{OPTIM}(P, A, B)$. We used the best solution found at a given resolution level as the initial mean vector for the first iteration at the next finer resolution level.
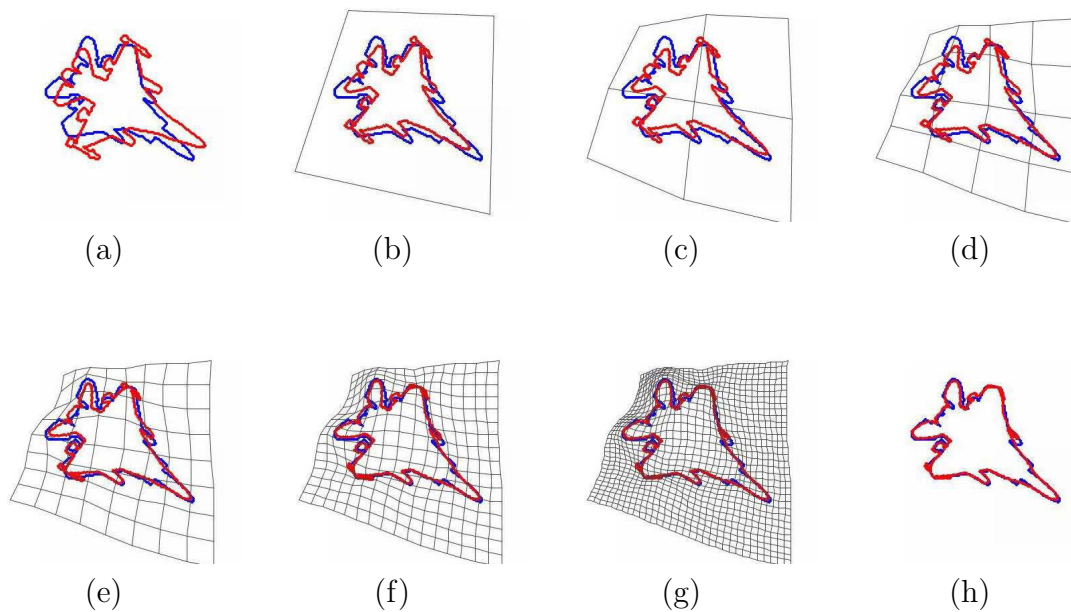
Figure 5.3: Multi-resolution FFD for shape registration. For the pairs of shapes in red and blue in (a) we show six multi-resolution levels ((b) - (g)), followed by the final registration in (h). At any given resolution level, a refinement operation creates a finer grid that allows finer deformations.

## 5.3  Results

Figure 5.3 shows the advantages of using a FFD model, a local robust Hausdorff distance and a multi-resolution approach. Figure 5.3(a) shows the reference shape (blue) and the floating shape (red) to be registered. Figures 5.3(b)-5.3(h) show six multi-resolution levels. Each level gives the algorithm the opportunity to achieve fines deformations. The first level (5.3(a)) performs a global perspective transformation of the floating shape. Figure 5.3(h) shows the final result and we can evaluate visually the good quality of the registration. There is only a slight problem with one of the wing engines. It is difficult to objectively evaluate non-rigid registration methods and throughout the literature visual evaluations are common ([PRR03, HPM06]). A similar display for three additional shapes can be seen in Figure 5.4. The initial conditions are shown in Column 5.4(a), results at a coarse multi-resolution level in 5.4(b), at a finer level in 5.4(c) and the final registration result in 5.4(d). We can see how the algorithm successfully brings the shapes into a global alignment and then proceeds to perform local deformations. Our method is able to register rough corners as in the fish top fin in 5.4(c). The FFD model tends to have difficulty finding control points that mimic these sharp deformations but the local Hausdorff distance is able to favor those situations. In general, the results are extremely good with minor misregistrations at very fine local level. We performed a quantitative evaluation for five different shapes and summarize it in Table 5.1. We counted the number of pixels in the symmetric difference image between reference and registered shapes (We used the binary shapes and not only the outline). The average number of pixels was 245 for our method and was 42% smaller than the results using Huang's method ([HPM06]).

We use the Hausdorff distance instead of the chamfer distance because the rank
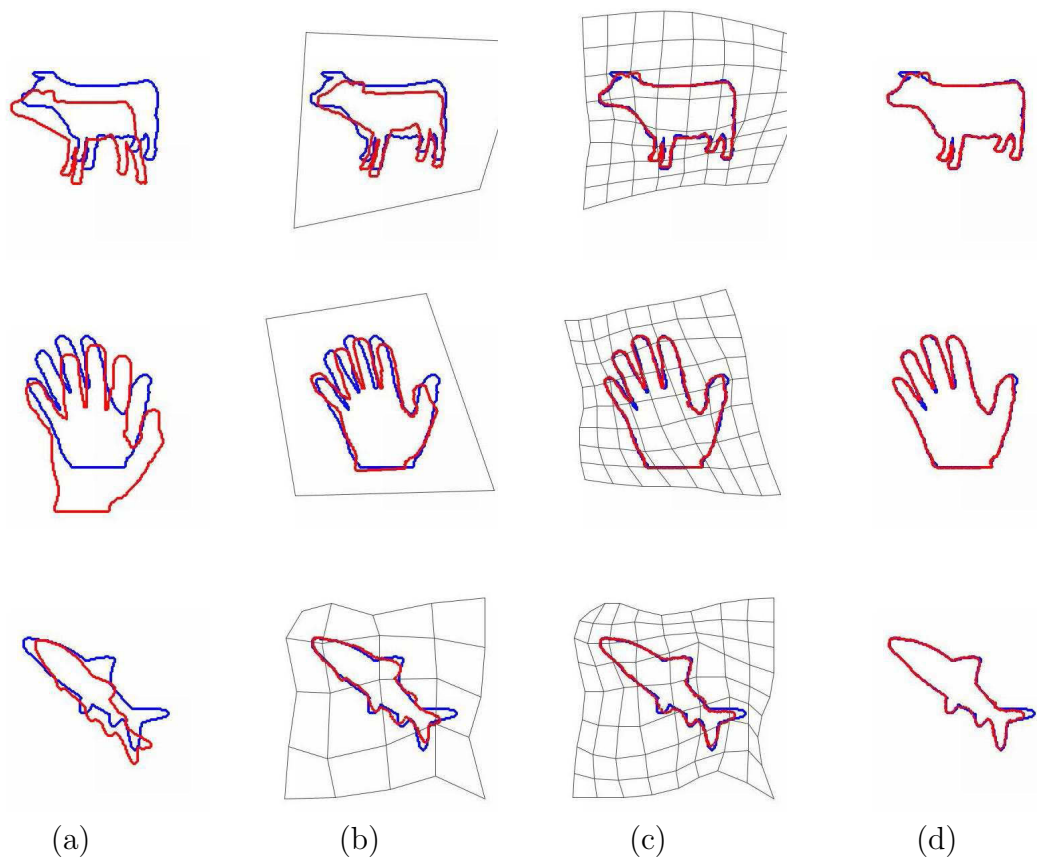
Figure 5.4: Hierarchical FFD point-based registration. For the pairs of shapes in red and blue in (a), we see the registration at the coarse level in (b), and at a finer level in (c), followed by the final registration in (d)

|  | Huang | Our Method |
|---|---|---|
|  | 348 | 162 |
|  | 611 | 380 |
|  | 708 | 219 |
|  | 354 | 195 |
|  | 118 | 271 |
| Average | 428 | 245 |

Table 5.1: Symmetric difference error. Pixel count in the image of the symmetric difference between reference and registered shapes

operation is robust to outliers while the average is still affected if a noise point is included. Both chamfer distance and Hausdorff distance tell only half of the registration story. They evaluate how close the points get to the zero level transform but they cannot tell which of those zero level DT values do not have a graph point close by. This leads to partial answers.

To compute the classical Hausdorff instead of the partial (directed) Hausdorff requires the computation of a distance transform or a Kd-tree for each intermediate transformation which is computationally expensive. For such cases, it might be useful to use the chamfer distance transform which is an approximation of the Euclidean distance.

Finding the maximum similarity value in the transformation space, whose dimensionality equals the degrees of freedom of the transformation, is a high dimensional optimization problem. This optimization is performed using iterative methods, which are generally sensitive to starting conditions and can get trapped in local minima.

### 5.3.1 Robustness to local minima

Figure 5.5 shows an example of shape registration where point-based methods get trapped easily in sub-optimal solutions. Huang et al. show that a TPS-ICP approach gets trapped because the misleading nearest neighbor computations prevent it from finding the right correspondences [HPM06]. Huang et al. also show that Context Shapes fails to differentiate local responses from hands and feet and creates false correspondences. Figure 5.5 shows how our algorithm is able to overcome that local minima. We show the registration results after using a 5x5 control point grid, which continues to have a poor registration. After a refining operation produces a 9x9 grid,
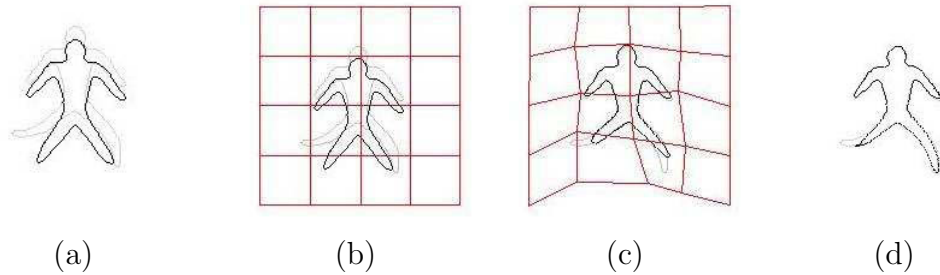
Figure 5.5: Starting point at local minimum

we see that the algorithm is able to register more accurately; right leg is registered well, but the algorithm fails to completely align the left leg (This shows a potential area of improvement where maybe local descriptors may be used at a local level.) Huang presents a visually optimal result for this case using a variational approach at the cost of maintaining a dynamic DT representation and computing an expensive estimation of the derivative of the energy functional with respect to the B-spline control grid parameters. We show satisfactory results using the Hausdorff distance which is not computationally expensive.

Another example of the robustness to local minima of our method is shown in Figure 5.6 where our method finds a good approximation to a hard problem with potential local minima regions. Although we do not achieve the optimal registration found by Huang et al. [HPM06], we compare visually and conclude that we achieve a better result for this experiment than what was achieved using the method by Paragios et al. [PRR03].

## 5.3.2   From shapes to images

We also show how our registration method for shapes can also be applied to registering 2D medical images. We tested the method by generating a simulated image

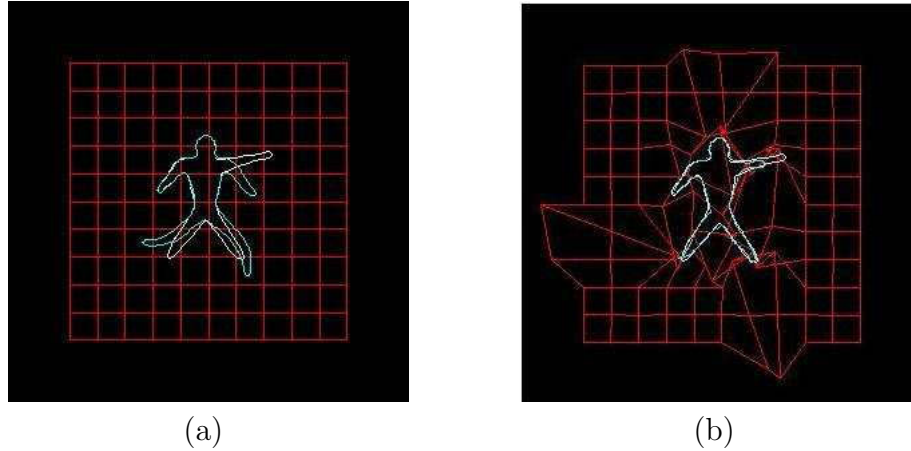<center>(a)                    (b)</center>

<center>Figure 5.6: Robustness to local minima</center>

distorted by applying a polynomial transformation (See Figure 5.7). We extracted a point set from each image, as shown overlapped in Figure 5.7(b). Then proceeded to compute the B-spline lattice control point locations that minimized the distance between the point sets and applied that function to the target image to map it as close as possible to the reference image. We used mutual information to measure registration quality. Before registration the MI was 0.47 and after registration 1.31. Figure 5.7 shows the result of our registration. The difference image (Figure 5.7(e)) has an intensity of 200 and is reduced to 20 in the difference of registered images (Figure 5.7(f)).

## Comments on the multi-resolution approach for images

Figure 5.8 shows the leverage of our method on the coarse-to-fine multi-resolution approach. Because of the term in Equation 5.3 that penalizes large node displacements in favor of local registrations, corresponding areas are required to be within the support area of the control point. In other words, if the corresponding point to be register is located outside of that support area, the algorithm is not going to find that correspondence. We solve this problem by relying on a multi-resolution

<center>118</center>

approach. It is good to remark that in a 3x3 grid all points within that grid can be reached from any other point. In Figure 5.8 we show the contrast between starting initially at a fine grid versus following the multi-resolution approach. It is clear from Figure 5.8(a) that the points from the part of the skull that suffer the most deformation fail to establish a correspondence when a fine grid is used, resulting in a suboptimal solution. Figure 5.8(b) shows the improved registration using the multi-resolution approach starting at a coarser lattice. This is also helpful because at finer resolution levels there is an increasing number of control points with empty support that cannot be used for registration introducing artifacts in the registration. Starting at a coarse level (a 3x3 grid has no points with empty support) the global registration helps to define the behavior for areas in finer lattices that would have empty support.

## 5.4 Conclusion

We present a robust point-based method that implicitly performs a global-to-local registration using a hierarchical free form deformation model that inherently provides smooth solutions. The presented method uses convenient embedded and graph point representations and is shown to be robust to noise and large local deformations (See Figure 5.8). Unlike other point-based methods (ICP, TPS), our method is also robust to starting conditions (Figures 5.5 and 5.6).

This novel point-based method uses a local Hausdorff distance measure and a B-spline based free from deformation opens the door for additional research to improve the local assessment of registration using local descriptors.

Our method uses the coarsest level (2x2) to compute a global registration. In 2D, this involves solving a problem with 8 degrees of freedom. Future work includes

computing a fast global transformation using points that estimate a rigid or affine initial approximation.

Future work also includes finding a mathematical relationship between our approach and a variational approach where the derivative of an energy functional with respect to the transformation parameters (control point location) is computed to guide the approximation.
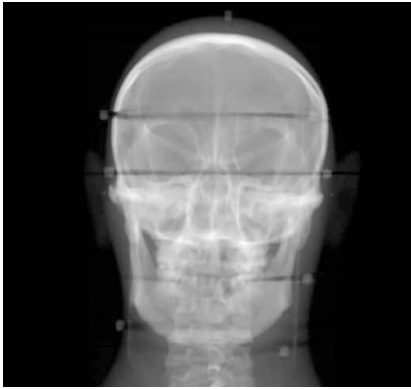
(a) reference image
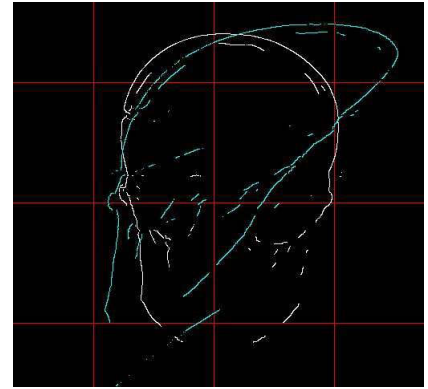


(b) point sets before registration
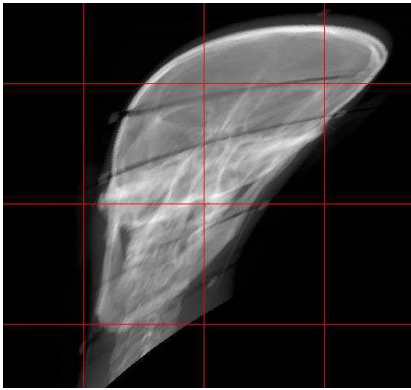


(c) target image
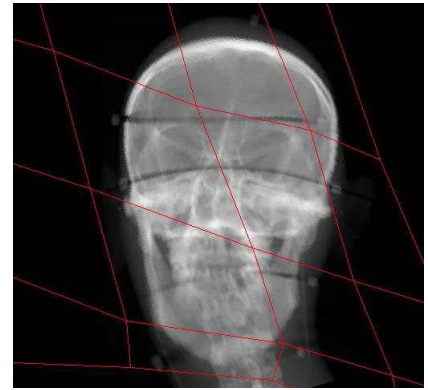


(d) registered target image



(e) difference image before registration



(f) difference image after registration

Figure 5.7: Point-based registration for image registration: We tested out point-based method to register a simulated image distorted by applying a polynomial transformation.
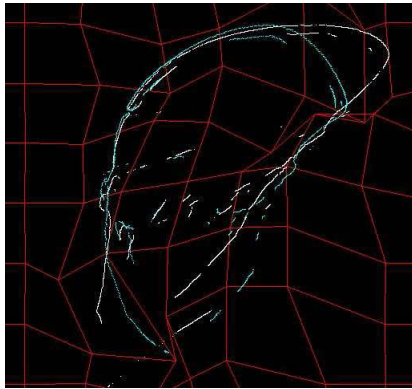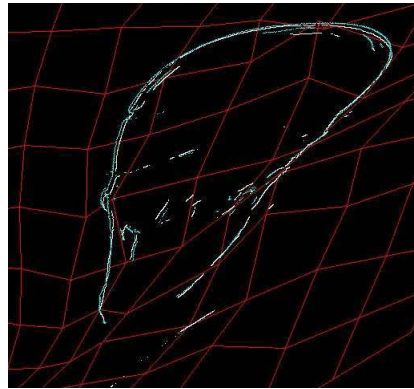
(a) without hierarchical FFD      (b) without hierarchical FFD

Figure 5.8: Registration using local Hausdorff distance

CHAPTER 6

## CONCLUSION

We conclude that our experimental results support our main claim that point-based registration of medical images can have performances that is comparable to or better than other existing methods. We achieve this by exploring similarity measures and search strategies that make up a robust framework for registration. In this chapter we review the general contributions of this dissertation and summarize how the different algorithms have helped us to achieve the goals outlined in the introductory chapter.

## 6.1   Contributions

The main goals of this dissertation were to explore point-based registration methods and to make them more accurate than existing (both point- and voxel-based) methods. Besides achieving these goals, in this dissertation we have developed a registration framework that is (a) robust, (b) can be extended to include local-descriptor information and (c) can be extended to non-rigid registration. For the application of intra-modal rigid registration in radiotherapy, the point-based method is more robust and more accurate than mutual information, the current leading method. For the non-rigid case, we have also shown that point-based methods can be made robust to starting conditions and that a local similarity measure allows accurate registrations for shapes.

In the case of rigid registrations, our central hypothesis is supported by proposing a robust hybrid search strategy that is shown to be better than its individual components, namely, the branch-and-bound strategy and the hill climbing approach, and also by extending the Hausdorff distance to be used in multi-modal and non-rigid registration. We conclude that point-based registration provides better results for

rigid registration of images in radiotherapy. We show that for EPID images, PoIRe is more robust than MI based methods and ICP. Finally, for non-rigid registration of shapes, we show that point-based methods perform better than competing methods.

All the above algorithms support *Goal 1,* which aims to explore point-based image registration methods and make them more accurate than existing methods. This is shown throughout the dissertation for intra-modal (Chapter 3), multi-modal (Chapter 4) and non-rigid registrations (Chapter 5). In Chapters 3-5 we consistently show how point-based registration methods provide accurate results in imaging for radiotherapy. None of the methods presented required the correspondences between pairs of points, thus satisfying *Goal 2*.

In Chapter 3, we showed that point-based registration methods can be more accurate than competing methods in applications that require the registration of intra-modal medical images which are very noisy (e.g., EPID images).

In Chapter 4, we established that PoIRe is more accurate than MI for radiotherapy images. This holds when the extracted points are associated with the response to a series of Gabor filters, which act as local descriptors (*Goal 3*). This chapter also presents a framework for using other local descriptors. In this chapter we also demonstrated that PoIRe produce accurate registration for multi-modal images (*Goal 4*)

Chapter 5 presents our method for non-rigid registration. Here we not only demonstrate experimentally the suitability of point-based registration for non-rigid registration (*Goal 5*), but successfully establish the superiority of our method to other (point-based and voxel-based methods).

## 6.2 Future work

We believe that there is a lot of potential for new directions mainly in the areas of point-based multi-modal image registration and in the case of non-rigid image registration.

### 6.2.1 Multi-modality registration

The first direction is to extend our robust 2D-2D registration to a 2D-3D scenario, where a collection of DRRs are generated from a reference CT image. The following direction is to include information from other local descriptors. Also, it would be useful to extend the Gabor functions to generate responses that are more orthogonal (i.e., less correlated).

### 6.2.2 Non-rigid registration

Here, we focused on the registration of shapes and only showed select examples on how to extend our approach from shapes to medical images. This is an interesting and promising area of further research. Our preliminary experiments suggest that point-based non-rigid registrations using a free-form deformation model are smooth enough to be used as a pre-processing step of methods like optical flow. We could also potentially include information from the Gabor filters or other local descriptors as part of the non-rigid local dissimilarity function. We showed in Chapter 4 that this is easily done for rigid registrations. Perhaps it could be extended to non-rigid registrations.

# BIBLIOGRAPHY

[AFP00]     M. A. Audette, F. P. Ferrie, and T. M. Peters. An algorithmic overview of surface registration techniquesfor medical imaging. *Medical Image Analysis*, 4(3):201–217, 2000.

[AGP06]     S. Arivazhagan, L. Ganesan, and S. Padam Priyal. Texture classification using gabor wavelets based rotation invariant features. *Pattern Recogn. Lett.*, 27(16):1976–1982, 2006.

[AHB87]     K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9(5):698–700, 1987.

[Ata85]     M. J. Atallah. A matching problem in the plane. *J. Comput. Syst. Sci.*, 31(1):63–70, 1985.

[AWS00]     L. Alvarez, J. Weickert, and J. Sánchez. Reliable estimation of dense optical flow fields with large displacements. *Int. J. Comput. Vision*, 39(1):41–56, 2000.

[BDNB06]    A. Bishnu, S. Das, S. C. Nandy, and B. B. Bhattacharya. Simple algorithms for partial point set pattern matching under rigid motion. *Pattern Recogn.*, 39(9):1662–1671, 2006.

[Ben90]     J. L. Bentley. K-d trees for semidynamic point sets. *Proc. 6th Annual ACM Symposium Computational Geometry*, pages 187–197, 1990.

[BJ88]      I. Biederman and G. Ju. Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 20:38–64, 1988.

[BM92]      P. J. Besl and N. D. Mckay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[BMP01]     S. Belongie, J. Malik, and J. Puzicha. Matching shapes. *Computer Vision, IEEE International Conference on*, 1:454, 2001.

[Boo89]     F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[Bor88]     G. Borgerfors. Chamfer matching: a parametric edge matching algorithm. *IEEE Transactions on pattern analysis and machine intelligence*, 10:849–865, 1988.

[BPC92]     M. Balter, C. A. Pelizzari, and G. T. Y. Chen. Correlation of projection radiographs in radiation-therapy using open curve segments and points. *Med. Phys.*, 19:329–334, 1992.

[Bro92]     L. G. Brown. A survey of image registration techniques. *ACM Comput. Surv.*, 24(4):325–376, 1992.

[Can86]     J. Canny. A computational approach to edge detection. *IEEE Trans on Pattern Anal Machine Intell*, 8:679–798, 1986.

[CSK05]     D. Chetverikov, D. Stepanov, and P. Krsek. Robust Euclidean alignment of 3D point sets: the trimmed iterative closest point algorithm. *Image and Vision Computing*, 23(3):299–309, 2005.

[Der93]     R. Deriche. Recursively implementing the gaussian and its derivatives. Technical Report 1893, INRIA, Unit de Recherche Sophia-Antipolis, 1993.

[DJ94]      M. P. Dubuisson and A. K. Jain. A modified hausdorff distance for object matching. volume 1, pages 566–568 vol.1, 1994.

[ELF97]     D. W. Eggert, A. Lorusso, and R. B. Fisher. Estimating 3-d rigid body transformations: a comparison of four major algorithms. *Mach. Vision Appl.*, 9(5-6):272–290, 1997.

[FA91]      W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:891–906, 1991.

[FB88]      D. R. Forsey and R. H. Bartels. Hierarchical b-spline refinement. In *SIGGRAPH '88: Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, pages 205–212, New York, NY, USA, 1988. ACM.

[FH04]      P. F. Felzenszwalb and D. P. Huttenlocher. Distance transforms of sampled functions. Technical report, Cornell Computing and Information Science, 2004.

[FWM98]    J. M. Fitzpatrick, J. B. West, and C. R. Maurer. Predicting error in rigid-body point-based registration. *Medical Imaging, IEEE Transactions on*, 17(5):694–702, 1998.

[GMO99]    M. T. Goodrich, J. S. B. Mitchell, and M. W. Orletsky. Approximate geometric pattern matching under rigid motions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):371–379, 1999.

[GMU96]    L. J. Van Gool, T. Moons, and D. Ungureanu. Affine/ photometric invariants for planar intensity patterns. In *ECCV '96: Proceedings of the 4th European Conference on Computer Vision-Volume I*, pages 642–651, London, UK, 1996. Springer-Verlag.

[GS99]     V. Di Gesú and V. Starovoitov. Distance-based functions for image comparison. *Pattern Recogn. Lett.*, 20(2):207–214, 1999.

[GTvHV95]  K. G. A. Gilhuijs, A. Touw, M. van Herk, and R. E. Vijlbrief. Optimization of automatic portal image analysis. *Med. Phys.*, 22:1089–1099, 1995.

[GvH93]    K. G. A. Gilhuijs and M. van Herk. Optimization of automatic portal image analysis. *Med. Phys.*, 20:667–677, 1993.

[HF96]     D. H. Hristov and B. G. Fallone. A grey-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs. *Med. Phys.*, 23:75–84, 1996.

[HM86]     A. Huertas and G. Medioni. Detection of intensity changes with subpixel accuracy using laplacian-gaussian masks. *IEEE Trans on Pattern Anal Machine Intell*, 8:651–664, 1986.

[Hor87]    B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.

[HPM06]    X. Huang, N. Paragios, and D. N. Metaxas. Shape registration in implicit spaces using information theory and free form deformations. *Transactions on Pattern Analysis and Machine Intelligence*, 28(8):1303–1318, 2006.

[HR93]     D. P. Huttenlocher and W. J. Rucklidge. 1993 a multi-resolution technique for comparing images using the hausdorff distance. *Proc*

*Third IEEE Conf Computer Vision and Pattern Recognition*, pages 705–706, 1993.

[JF91]        A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recogn.*, 24(12):1167–1186, 1991.

[Kan94]       K. Kanatani. Analysis of 3d rotation fitting. *IEEE PAMI 16*, 5:543–549, 1994.

[KFL$^+$01]   J. Kim, J. A. Fessler, K. L. Lam, J. M. Balter, and R. K. TenHaken. A feasibility study of mutual information based set-up error estimator for radiotherapy. *Med. Phys.*, 28:2507–2517, 2001.

[KKK02]       J. K. Kamarainen, V. Kyrki, and H. Kälviäinen. Fundamental frequency gabor filters for object recognition. In *ICPR (1)*, pages 628–631, 2002.

[KKK06]       J. K. Kamarainen, V. Kyrki, and H. Kälviäinen. Invariance properties of gabor filter-based features-overview and applications. *IEEE Transactions on Image Processing*, 15(5):1088–1099, 2006.

[KPP95]       V. Kvasnicka, M. Pelikan, and J. Pospichal. Hill climbing with learning (an abstraction of genetic algorithm). *Proceedings of the First International Conference on Genetic Algorithms. Brno, Czech Republic*, pages 65–70, 1995.

[KWT87]       M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, January 1987.

[LLB98]       K. W. Leszczynski, S. Loose, and S. Boyko. An image registration scheme applied to verification of radiation therapy. *Brit. J. Radiol.*, 71(884):413–426, 1998.

[LLS03]       K. H. Lin, K. M. Lam, and W. C. Siu. Spatially eigen-weighted hausdorff distances for human face recognition. *Pattern Recognition*, 36(8):1827–1834, 2003.

[LLS04]       D. H. Liu, K. M. Lam, and L. S. Shen. Optimal sampling of gabor features for face recognition. *Pattern Recogn. Lett.*, 25(2):267–276, 2004.

[Low99]      D. G. Lowe. Object recognition from local scale-invariant features. In *Proc. of the International Conference on Computer Vision ICCV, Corfu*, pages 1150–1157, 1999.

[LVM02]      J. Liu, B. C. Vemuri, and J. L. Marroquin. Local frequency representations for robust multi-modal image registration. *IEEE Transactions on Medical Imaging*, 21:462–469, 2002.

[Mar63]      D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.

[MCV+97]     F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging*, 16(2):187–198, April 1997.

[MF93]       C. R. Maurer and J. M. Fitzpatrick. A review of medical image registration, 1993.

[MFW+97]     C. R. Maurer, J. M. Fitzpatrick, M. Y. Wang, R. L. Galloway, R. J. Maciunas, and G. S. Allen. Registration of head volume images using implantable fiducial markers. *IEEE Trans. Med. Imag*, 16(4):447–462, 1997.

[MLS05]      K. Mikolajczyk, B. Leibe, and B. Schiele. Local features for object class recognition. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision*, pages 1792–1799, Washington, DC, USA, 2005. IEEE Computer Society.

[MMA06]      C. J. Moore, T. E. Marchant, and A. M. Amer. Cone beam ct with zonal filters for simultaneous dose reduction, improved target contrast and automated set-up in radiotherapy. *Phys. Med. Biol.*, 51(9):2191–2204, 2006.

[MNL98]      D. M. Mount, N. S. Netanyahu, and J. LeMoigne. Efficient algorithms for robust feature matching. *Pattern Matching*, 32:17–28, 1998.

[Mur02]      M. J. Murphy. Fiducial-based targeting accuracy for external-beam radiotherapy. *Med. Phys.*, 29:334–344, 2002.

[MV98]       J. Maintz and M. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2:1–36, 1998.

[MVS99]      F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3(4):373–386, 1999.

[PFY⁺99]     S. M. Pizer, D. S. Fritsch, P. A. Yushkevich, V. E. Johnson, and E. L. Chaney. Segmentation, registration, and measurement of shape variation via image object shape. *IEEE Transactions on Medical Imaging*, 18:851–865, 1999.

[Pow64]      M. J. D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, 7(2):155–162, February 1964.

[PP07]       N. A. Parra and C. A. Parra. Gabor feature-based registration: accurate alignment without fiducial markers. In *Proceedings of SPIE Medical Imaging*, volume 6512, 2007.

[PRR03]      N. Paragios, M. Rousson, and V. Ramesh. Non-rigid registration using distance functions. *Comput. Vis. Image Underst.*, 89(2-3):142–165, 2003.

[PST⁺00]     D. Plattard, M. Soret, J. Troccaz, P. Vassal, J. Y. Giraud, G. Champleboux, X. Artignan, and M. Bolla. Patient set-up using portal images: 2d/2d image registration using mutual information. *Computer Aided Surgery*, 5(4):219–306, 2000.

[RK96]       S. Rudlof and M. Koppen. Stochastic hill climbing with learning vector of normal distributions. *1st Online Workshop on Soft Computing, Retrieved from http://www.bioele.nuee.nagoya-u.ac.jp/wscl/*, 1996.

[SC06]       S. C. Xu N. L. Sun and M. Y. Cao. Wall-pasted cell segmentation based on gabor filter with parameter constraint. pages 1–6, Dec. 2006.

[Sch66]      P. Schnemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, March 1966.

[SD02]     D. Shen and C. Davatzikos. HAMMER: hierarchical attribute matching mechanism for elastic registration. *Medical Imaging, IEEE Transactions on*, 21(11):1421–1439, 2002.

[SHF99]    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. *Medical Image Analysis*, 26(11):2422–2428, 1999.

[SKP99]    D. G. Sim, O. K. Kwon, and R. H. Park. Object matching algorithm using robust hausdorff distance measures. *IEEE Trans. Image Processing*, 8:425–429, 1999.

[SP86]     T. W. Sederberg and S. R. Parry. Free-form deformation of solid geometric models. *SIGGRAPH Comput. Graph.*, 20(4):151–160, 1986.

[Tak98]    B. Takács. Comparing face images using the modified hausdorff distance. *Pattern Recognition*, 31(12):1873–1881, 1998.

[TU00]     P. Thevenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.

[TWK88]    D. Terzopoulos, A. Witkin, and M. Kass. Constraints on deformable models: recovering 3d shape and nongrid motion. *Artif. Intell.*, 36(1):91–123, 1988.

[Ume91]    S. Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(4):376–380, 1991.

[VW95]     P. Viola and W. M. Wells. Alignment by maximization of mutual information. *Fifth Intl. Conference on Computer Vision*, 1:16–23, 1995.

[WC05a]    Y. Wang and C. S. Chua. Face recognition from 2d and 3d images using 3d gabor filters. *Image Vision Comput.*, 23(11):1018–1028, 2005.

[WC05b]    Y. Wang and C. S. Chua. Face recognition from 2d and 3d images using 3d gabor filters. *Image and Vision Computing*, 23(11):1018 – 1028, 2005.

[WC06a]     Y. Wang and C. S. Chua. Robust face recognition from 2d and 3d images using structural hausdorff distance. *Image Vision Comput.*, 24(2):176–185, 2006.

[WC06b]     Y. Wang and C. S. Chua. Robust face recognition from 2d and 3d images using structural hausdorff distance. *Image Vision Comput.*, 24(2):176–185, 2006.

[WFKvdM97] L. Wiskott, J. M. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *Image Processing, International Conference on*, 1:129, 1997.

[WFW+97]    J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, T. S. Sumanaweera, B. Harkhess, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Mal, X. Pennec, M. E. Noz, G. Q. Maguire, M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods. Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography*, 21:554–566, 1997.

[WT03]      X. Wang and X. Tang. Bayesian face recognition using gabor features. In *WBMA '03: Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications*, pages 70–73, New York, NY, USA, 2003. ACM.

[YLC04]     C. H. T. Yang, S. H. Lai, and L. W. Chang. Robust face image-matching under illumination variations. *EURASIP J. Appl. Signal Process.*, 2004:2533–2543, 2004.

[YLC07]     C. H. T. Yang, S. H. Lai, and L. W. Chang. Hybrid image matching combining hausdorff distance with normalized gradient matching. *Pattern Recogn.*, 40(4):1173–1181, 2007.

[ZF03]      B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, October 2003.

[Zhu02]     Y. M. Zhu. Volume image registration by cross-entropy optimization. *Medical Imaging, IEEE Transactions on*, 21(2):174–180, Feb. 2002.

[ZTL04]     Z. Zhu, M. Tang, and H. Lu. A new robust circular gabor based ob-
            ject matching by using weighted hausdorff distance. *Pattern Recogn.
            Lett.*, 25(4):515–523, 2004.

[ZTM02]     J. Zhang, T. Tan, and L. Ma. Invariant texture segmentation via
            circular gabor filter. In *Proceedings of the 16th IAPR International
            Conference on Pattern Recognition (ICPR)*, Vol II, pages 901–904,
            2002.

[ZWIL00]    D. S. Zhang, A. Wong, M. Indrawan, and G. Lu. Content-based
            image retrieval using gabor texture features. In *IEEE Transactions
            PAMI*, pages 13–15, 2000.

# VITA

## NESTOR ANDRES PARRA

## AREAS OF INTEREST

Medical image processing, computational geometry, object-oriented software development, evolutionary computation, internet programming.

## EDUCATION

| | |
|---|---|
| 2009 | Doctoral Candidate, Computer Science<br>School of Computing and Information Sciences<br>Florida International University<br>Miami, Florida |
| 2002 | M.S., Mathematical Science<br>The University of Memphis<br>Memphis, Tennessee |
| 1998 | B.S., Computer and System Engineering<br>National University of Colombia<br>Bogota, Colombia |

## PUBLICATIONS AND PRESENTATIONS

Parra N. A. and Parra C. A., (2007). *Gabor feature-based registration: accurate alignment without fiducial markers.* In Proceedings of SPIE Medical Imaging, 1(6512).

Parra A., Narasimhan G., and Samant S., (2005). *A robust marker registration algorithm.* Medical Physics, 32(6).

Samant S. S., Zheng W., Parra N. A., Chandler J., Gopal A., Wu J., Jain J., Zhu Y., and Sontag M, (2002). *Verification of multileaf collimator leaf positions using an electronic portal imaging device.* Medical Physics, 29(12).

Samant S., Parra N., Davis B. C., Sontag M., and Narasimhan G. A new multimodality image registration algorithm, (2002). *Gabor feature-based registration: accurate alignment without fiducial markers.* In Proceedings of the 44th Annual Meeting of the American Association of Physicists in Medicine (AAPM), 1(3216).

Nino F., Hernandez G., and Parra A., (2000). *Financial time series modeling with evolutionary trained random iterated neural networks.* In Proceedings of the Conference on Computational Intelligence for Financial Engineering (CIFER).