Point-based nonrigid registration : application to object recognition and medical image registration

Ahmad, Fahad Hameed

2012

Ahmad, F. H. (2012). Point-based nonrigid registration : application to object recognition and medical image registration. Master' s thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/48205>

<https://doi.org/10.32657/10356/48205>

Point-based Nonrigid Registration: Application to Object Recognition and Medical Image Registration

Fahad Hameed Ahmad

School of Computer Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Master of Engineering

2012

Acknowledgements

First and foremost I offer my sincerest gratitude to my supervisor, Dr Sudha Natarajan, who has supported me throughout my thesis with her patience and knowledge whilst allowing me to work in my own way. I would also like to thank my co-supervisor Dr Jimmy Liu Jiang and all my colleagues at I2R who gave me sound advice and helped me in acquiring resources.

Lastly, I offer my regards to my covering supervisor Dr. Suresh Sundaram and blessings to all of those who supported me in any respect during the completion of this thesis.

Fahad Hameed Ahmad

Abstract

Image registration is an important problem in computer vision and has many diverse applications. Registration is the method of aligning two or more images into the same coordinate system to achieve a one-to-one correspondence. While registration of images has been studied in the past, an area that is less explored is nonrigid registration. It turns out that nonrigid registration is very appropriate when complex deformations are involved. In this thesis, we focus on point-based nonrigid registration and develop algorithms for solving nonrigid registration problems.

With nonrigid registration as the goal, new point-based shape descriptor for object classification has been developed. The shape descriptor is based on principal curvatures and their directions. High curvature boundary points are first extracted in a multi-scale environment. The shape descriptor is then extracted by taking the k nearest neighborhood boundary points around a point of interest and calculating the central moments of the distribution of different properties of these neighborhood points. The choice of scale and the number of nearest neighbors provides a global as well as a local description of the neighborhood. Compared to previous techniques for shape matching and classification, experiments show that the proposed algorithm is more robust to deformation and invariant to rotation, scaling and translation.

We also present a robust registration method based on geometric invariant features using multi-scale weighted quaternion sub-division. The geometric invariant features are based on selective crest points which are robust to deformation. The k -nearest neighbor approach combined with multi-scale provides a new adaptive method for collecting features. The size of the neighborhood is increased if fewer features are present and is decreased when features are abundant. A feature of quaternion fitting approach is that it provides a fast method for registering points during sub-division registration, and the associated feature weights help remove any outlier effect on the result of the registration. A rigid body deformable registration scheme is also presented which is robust to noise, intensity variation and missing features. Hausdorff distance has been used as a similarity measure. The registration technique focuses on multi-modal medical images which have significant intensity variations and different levels of SNR. A huge amount of feature mismatch is also present between modalities as different modalities depict different tissues differently. We have analyzed the robustness and sensitivity of our algorithm with respect to various CT and MRI images of different quality.

Contents

Acknowledgements	i
Abstract	ii
1 Introduction	1
1.1 Research Objectives	6
1.2 Major Contributions	7
1.3 Publications	8
1.4 Outline	8
2 Literature Review	9
2.1 Image Registration	9
2.1.1 Intensity-based Registration	10
2.1.2 Feature-based Registration	12
2.2 Shape Descriptors for Shape Recognition, Retrieval and Classification	15
2.2.1 Region-based Feature Descriptor	16
2.2.2 Contour-based Feature Descriptors	16
2.2.3 Shape Matching using Similarity Measures	18
2.3 Conclusion	19
3 New Shape Descriptor for Object Classification	20
3.1 Proposed Shape Descriptor	21
3.1.1 Extracting High Curvature Interest Points	21
3.1.2 Feature Extraction	25
3.1.3 Moment-based Shape Descriptors	28
3.1.4 Feature Normalization	29
3.2 Object Classification	29
3.2.1 Image Alignment using Matched Points	31
3.2.2 Outlier Removal and Iteration	31
3.2.3 Cost Function	31
3.3 Conclusion	32
4 Performance Studies and Comparisons	33
4.1 Choice of Various Parameters for Optimizing Shape Matching	33
4.1.1 Choice of "k" for Nearest Neighbor Search for Feature Extraction	33
4.1.2 Choice of Number of Scales σ	34

4.2	Studies on Different Databases	34
4.2.1	Sclaroff Tool Database	34
4.2.2	Kimia-99 database	35
4.2.3	Kimia-25 Database	38
4.2.4	Articulated Tool Data Set	38
4.2.5	MNIST Handwritten Digit Database	39
4.3	Computation Time	41
4.4	Conclusion	41
5	Point-based Non-rigid Registration of Medical Images	43
5.1	Deformable Rigid Body Hausdorff Registration for Multi-modal Medical Images	43
5.1.1	Differential Geometry-based Feature Extraction for Image Registration	43
5.1.1.1	Differential Geometry - An Overview	44
5.1.1.2	Finding Crest Points through Principal Curvatures	44
5.1.1.3	Multi-scale Analysis and Feature Extraction	45
5.1.2	3D Rigid Body Registration-based on the Hausdorff Similarity Measure	47
5.1.2.1	Multi-scale Feature Sub-division	47
5.1.2.2	Conventional Hausdorff Distance	48
5.1.2.3	Optimization	49
5.1.3	Interpolation and Reconstruction - Deformation Field	50
5.1.4	Experiments	50
5.2	Feature Based Nonrigid Registration using Quaternion Sub-division	52
5.2.1	Feature Point Extraction	53
5.2.1.1	Geometric Invariant Feature Extraction	54
5.2.2	Registration Using Quaternions	56
5.2.2.1	Sub-division Registration Using Unit Quaternion	58
5.2.3	Deformation Field and Interpolation	58
5.2.4	Experiments	59
5.3	Conclusion	59
6	Conclusion and Future Work	62
6.1	Summary	62
6.2	Possible Extensions and Future Work	63
A	Geometric Primitives	66
B	Low Order Moments	67
C	Quaternion Fitting	68
	Bibliography	70

Chapter 1

Introduction

Image registration is an important task in computer vision. It has many applications in medical image processing, shape recognition and stereo matching. Some medical applications for registration include aligning multi-modal images, Image Guided Interventions (IGI) to match pre- and post-operative images, probability atlas construction for segmentation, and comparing inter-patient data in the course of a treatment to compensate for breathing and other movements. In stereo matching, landmark points are generally matched to estimate the depth of the objects present in the scene. Similarly, point correspondences could be used for matching and classifying objects.

Registration is loosely defined as a transform that maps points from one image to corresponding points in another image based on a similarity metric. These data points can be extracted from different modalities, different time instants or from different subjects. The image that undergoes transformation during registration is called the moving image, whereas the other image is referred to as the fixed image.

Registration is further divided into two categories: rigid and nonrigid registration. In medical applications, rigid registration is useful in registering intra-patient images that undergo very little deformation with time. The shape and size of a human brain varies very little and only the head position changes, which makes rigid registration suitable for registering brain MRI scans. MRI brain scans are shown in Fig. 1.1 where the rigidity of the cranial structure is visible. On the other hand, flexible organs in the abdomen change both shape and size which requires nonrigid registration. Nonrigid registration is a challenging problem and has not been studied much in the literature. Fig. 1.2 shows inter-patient data for two patients. It can be seen from the figure that internal organs in the abdomen are flexible and are capable of changing their positions.

FIGURE 1.1: Two different brain MRI (T1 and T2) images. T1 and T2 weighted scan highlight different regions of the brain and an overlap between the different brain areas would help doctors in their evaluation

FIGURE 1.2: Abdomen CT scans are shown in the figure for two different patients. Comparison between different patients is at times required to identify abnormalities. The gall bladder in the abdomen CT is marked with a red boundary and a change in shape and size could be observed.

A detailed survey on different nonrigid registration algorithms is provided in [1]. Registration methods are broadly categorized into two: feature-based registration and intensity-based registration. Feature-based registration methods identify salient features like points, lines or surfaces in the moving image and fixed images before finding the corresponding features and transformation. On the other hand, intensity-based registration methods can be applied directly to the images without any feature extraction step. Feature-based methods use salient parts of the image to drive the registration and once a transformation is established then that transformation is

FIGURE 1.3: Noise due to pixel rotation is shown, where the original image on the left is rotated by twenty-five degrees. The results for bicubic (center) and nearest neighbor (right) interpolation are shown.

FIGURE 1.4: Original and rotated points are shown. Note that there is no constraint for rotating points sets. The original extracted points are shown on the left and the rotated points sets are shown on the right

interpolated over the entire image. Since points are represented in coordinate space, they can take on non-integer values which simplifies handling points or contours. On the other hand, intensity-based deformation measures are less flexible as it requires interpolation to be carried out at every step as not every location is an integer. Fig. 1.3 shows the artifacts produced in rotating and interpolating pixels, where as Fig. 1.4 shows no such problems. Mutual information is one popular intensity-based method. Mutual information is unstable for out of plane rotation [2] and fails due to lower resolution when the images contain little information, small region overlap or as a result of interpolation methods [2].

In image-guided surgeries, an intraoperative image has to be registered with a preoperative image. The images can differ in their modality. Different types of medical images provide different details of the same underlying anatomy and present the details in varying color or intensity. Medical imaging techniques can be broadly classified into two categories, functional and anatomical. MRI, CT and X-ray are often used for anatomical imaging. In addition to these techniques, many video sequences are also obtained using scopes and probes, for example in laproscopic procedures, video sequences need to be matched with pre-operative CT and MRI scans for surgical planning. Functional imaging techniques generally highlight the functioning of the anatomy. SPECT(Single Photon Emission Computed Tomography), PET(Positron Emission Tomograph) and fMRI(functional MRI) are some of the other widely used techniques [3]. Similarly, real world images of objects vary in color and intensity based on the lighting conditions. Many heuristics based approaches are available which use color as a feature for describing regions [4] but these are not robust to variations in the environment and are not easily generalizable.

Lessons learnt from cognitive science suggest that edges and points are more robust and useful for identifying objects. Edge features are used by humans for visual recognition [5] and have been shown to be better at identifying objects than color or intensity [6].

FIGURE 1.5: Two examples of (a) rotational symmetry and (b) reflective symmetry are shown. The point descriptors formulated for finding point correspondences should be able to differentiate between symmetric points.

Point matching and registration is also an important step in stereo matching and object recognition [7]. Corresponding points are found between similar images and alignment is done to find a suitable measure for the closeness of matched points. Point descriptors should not only be rotation, translation and scale invariant but should also be able to cope with symmetric objects. Symmetry is often hard to differentiate but many global measures like principal axis or geometric moments could be used to differentiate between reflective symmetry as well as rotational symmetry [8]. Fig. 1.5 shows two cases of rotational and reflective symmetry. The descriptor formulated should also be robust to noise, occlusion and articulation. Occlusions and articulation are a frequent occurrence in real-world objects.

Points are the most basic building blocks of edges and surfaces. Point-based registration methods are faster than intensity-based registration methods as fewer points are required to calculate the alignment [9]. Many closed form solutions are available for finding the alignment once point correspondences are found [10]. One drawback of point-based registration scheme is that it requires some form of point descriptor to find corresponding points. High dimensional point descriptors are generally used for matching which could be computationally expensive but would lead to a simpler and faster transformation and alignment step. Feature-based approaches using points, lines or surfaces perform better than intensity-based methods for multi-modal registration in medical images.

Fig. 1.6 illustrates point-based registration. It is roughly divided into two steps. The first step assigns point correspondence while the second step finds a suitable transformation based on the correspondence assigned. Many variants exist for this rough model, some search for point correspondence explicitly while others focus on finding a suitable transformation without explicitly

FIGURE 1.6: Block Diagram for Point-based Registration

assigning correspondences. Principal axis, center of mass and moments could be used to find a suitable transformation without the need for finding point correspondence [11]. Similarly, similarity measures like the Hausdorff distance could be minimized to find the best transformation [12]. These methods cannot be extended easily to nonrigid registration of points and lines due to the higher complexity and nonlinearity of the problem. Nonrigid transformations are easier to model once point correspondences are found. The most popular algorithm is the ICP [13] which uses k-means to assign crude correspondences and then refines the transformation to converge at a local minimum in a least squares sense. In case of nonrigid registration, this algorithm degrades quickly and is not well suited if too many outliers, occlusions or articulations are present which makes it sensitive to the initial point correspondences assigned. Another popular method is to formulate point descriptors and focus on finding reliable point correspondences before moving to the transformation step. One popular category of point descriptors are derived from the work of Belongie et al. [14]. Point descriptors also simplify the problem into bipartite matching compared to the more difficult NP-complete graph matching problem. Transformations obtained using point descriptors are more robust to large deformations.

The discussion so far indicates that nonrigid registration has not been thoroughly studied and there is scope for enhancement particularly from the point of view of accuracy and speed [9, 15]. Anatomical and functional correspondences are hard to define mathematically and there is room for much improvement. The need for registration extends to many different modalities which requires a more holistic approach and a need to move away from the more popular intensity-based methods.

1.1 Research Objectives

- **Goal 1: Explore point-based nonrigid registration scheme.** Nonrigid registration is a challenging problem with high computational complexity. The large number of transformation parameters makes it susceptible to outliers and noise and leads to an ill-posed problem. A practical nonrigid point set registration algorithm should have high accuracy and tractable complexity. One way to reduce the complexity is to have higher dimensional feature descriptors to find reliable point correspondences followed by a simpler transformation with far fewer parameters based on those point correspondences. A higher dimensional feature descriptor would make point set correspondences less likely to degrade in the presence of noise and outliers.
- **Goal 2: Develop point-based descriptor for better estimation of point correspondences.** Point features are generally extracted using underlying image properties. Unlike intensity-based matching and registration which takes the entire image into consideration, point-based matching focuses on points with salient and differentiable properties which should be matched. Most voxel-based registration algorithms focus on registering regions which would decrease the accuracy of specific areas which may have more importance compared to regions with sparse features. Many nonrigid registration cannot perfectly match and register every voxel due to the underlying differences in the imaging techniques (as in the case of multimodal images). Point-based matching and registration therefore is inherently superior due to its focus on registering important and salient points only. Most point-based registration schemes require extraction of geometric features and matching step needs to be spatially constrained. There is therefore a need to develop better feature descriptors which would be able to handle the local and global nonlinear deformations.
- **Goal 3: Develop fast transformation algorithms** Fast transformation algorithms are developed for registering objects nonrigidly once point correspondences are found. One approach followed was to develop rigid body subdivision schemes. A lot of research has already been done on developing fast transformation algorithms for rigid registration and this work could be extended to nonrigid registration by using subdivision schemes.

1.2 Major Contributions

The primary contributions of this thesis are as follows. We first present a novel shape descriptor based on principal curvatures and their directions. This is valuable in object recognition and classification. High curvature boundary points are first extracted in a multi-scale environment. The shape descriptor is extracted by taking the k nearest neighborhood boundary points around an interest point and calculating the central moments of the distribution of different properties of these neighborhood points. The choice of scale and the number of nearest neighbors provide a global as well as a local description of the neighborhood. Once point correspondences between images are found, a simple iterative closest point algorithm is used to align the boundary points of two different shapes. Euclidean distance is then computed to measure the similarity between matching points after alignment, both in the coordinate space and in the feature space for object recognition and classification. Compared to previous techniques for shape matching and classification, experiments show that the proposed algorithm is more robust to deformation and invariant to rotation, scale and translation.

The next contribution is a new deformable rigid body technique which uses differential geometry based features and the Hausdorff similarity measure for rigid body sub-division registration. The differential geometry based features and the Hausdorff similarity measure are chosen because of their robustness to missing features and intensity variations. A coarse to fine optimization was performed by constructing a multi-scale pyramid and then dividing the image into sub-templates as resolution was increased at finer scales. A translation deformation field is obtained which maps each pixel to a new location followed by a cubic interpolator to get the registered image. Experimental results show that the new registration technique offers better accuracy and quality, and is computationally less intensive compared to previous methods.

Another contribution is a new weighted quaternion based nonrigid registration scheme. Strong crest points derived from principal curvatures provide the most robust features for image registration. Crest point weights are based on their principal curvatures and the number of scales a particular crest point is detected at. Geometric features are extracted which are invariant to rotation, translation and scaling by using neighborhood crest points only as other voxels are susceptible to deformation. The neighborhood size is adjusted according to scale adaptively using a fixed k nearest neighbor to make the extracted feature scale invariant. Statistical properties are used to measure the distribution of these geometric invariant features. The scale and rotation invariant feature points are then used to establish a point to point correspondence between the template crest points and the subject image crest points. A multi-scale feature based subdivision scheme is employed for registration where a weighted quaternion matrix provides a quaternion transformation based on the corresponding points to obtain the best rotation for global as well as local sub-blocks.

1.3 Publications

1. F. H. Ahmad and S. Natarajan, “Multi-modal Medical Images Registration Using Differential Geometry and Hausdorff Distance”, *Journal of Intelligent Systems*, Vol. 19, No.4, 2010.
2. Fahad Hameed Ahmad, N. Sudha, Jimmy Liu Jiang, “Feature based nonrigid registration using quaternion subdivision”, *Proceedings of International Conference on Medical Biometrics* (LNCS 6165), Hong Kong, June 2010.
3. Fahad Hameed Ahmad, N. Sudha, “Deformable rigid body Hausdorff registration for multi-modal medical images”, *Proceedings of IEEE TENCON*, Singapore, Nov. 2009.
4. Fahad Hameed Ahmad, N. Sudha, “Feature-based Directional Shape Descriptor for Shape Retrieval and Classification”, *Machine Vision and Applications* (under review).

1.4 Outline

In Chapter 2, we present a brief literature survey on image registration. Merits and demerits of both intensity-based and feature-based registration methods are discussed and an overview is presented on the constraints associated with nonrigid registration is presented. In Chapter 3 we introduce a new point descriptor for shape recognition and object retrieval. Experimental results and comparison studies on different image databases are given in Chapter 4 to test the performance of the new point-based descriptor introduced earlier. In Chapter 5, we present an extension of proposed point-based descriptor applicable to medical image registration. This chapter is divided into two sections; Hausdorff distance sub-division registration is presented in one section while the other section describes a quaternion weighted sub-division scheme along with some experimental results. Chapter 6 summarizes the contributions and suggests future directions.

Chapter 2

Literature Review

In this chapter, we give a detailed review of literature in image registration especially for medical images and also present various point-based descriptors for shape retrieval and classification.

2.1 Image Registration

Image registration is an important task in many image processing applications such as stereo matching [16], image retrieval [17], shape recognition [14] and medical applications [18]. Alignment, fusion, matching, stabilization or stitching are all different names for registration based on the application.

Registration has found widespread use in medical imaging where the large number of imaging modalities and clinical and diagnostic procedures creates a need for a common reference. Some of the common applications for medical image registration include alignment of data sets belonging to different modalities [19] [20], aligning pre- and post-scan data sets for diagnostic purposes [21], [22], registering different modalities for surgical guidance [23], [24], atlas-based segmentation of different anatomical and functional regions [25], [26], [27], [28], [29], [30], [31], creating and using models for understanding the human anatomy [32], and for classification purposes [33], [34].

Different modalities exist for different diagnostic procedures. Registration of time series images in fMRI studies facilitates the detection of any changes in brain function and anatomy [35]. It also helps in correcting patient motion by registering the time series images to an anatomical image of the brain [36]. Similarly, CT registration is helpful to validate procedures like pedicle screw placement or root canal treatment [37], [38]. Ultrasound registration is frequently used in breast treatment [39] and to monitor cardiac functions [40]. Registration has also been employed in analyzing retinal fundus images [41]. Multi-modality registration is often necessary to get

more detailed information about a problem area. CT and MR images are frequently registered to improve diagnosis [42]. In total, there are many different combinations and strategies for applying and using registration in medicine.

In image registration, one image $I_F(x)$ is fixed which undergoes no deformation, while the second image is the moving image $I_M(x)$ which is transformed to fit the coordinate space of the fixed image. The aim is to find this transformation $T(x)$ which would give us the best transformed data set $I_M(T(x))$ such that some cost function $C(T; I_F, I_M)$ is minimized.

Registration itself is an ill-posed problem, since many registration problems are not defined properly. According to Hadamard, a well posed problem is one whose solutions exists, has a unique solution and is continuous. Registering inter-patient data can never be truly possible due to the inherent differences in the source of the data sets. The data sets are similar in some ways but also differ widely, hence the definitions of similarity and cost function $C(T; I_F, I_M)$ varies which creates a need for a regularization framework that penalizes undesirable properties. In some applications, rigid registration may be sufficient, and the degrees of freedom for $T(x)$ can be reduced, but in other cases where non-linear deformations are present, it would be important to increase the degrees of freedom. Most of the human body does not conform to a rigid or even an affine transformation and registration problem can only be solved in such a case using nonrigid means. In this report, we would focus only on nonrigid registration which is more generic and if efficient and fast nonrigid algorithms are developed then these algorithms could altogether replace rigid registration algorithms as well.

Most registration algorithms developed focus on targeted applications. Industrial applications focus on the speed of the algorithms, registering high resolution images needs high memory requirements, other applications focus on the accuracy and speed of the algorithm which may be required in surgical planning. [43].

Registration algorithms are broadly divided into two categories 1) Intensity-based and 2) Feature based [18], [44].

2.1.1 Intensity-based Registration

Intensity-based registration operates directly on the images without a feature extraction step, seeking to minimize some intensity-based similarity measure. [45], [46], [47], [48]. Some common intensity based similarity metrics or cost functions are Sum of Squared Differences (SSD), Correlation Coefficient (CC), Correlation Ratio (η), Mutual Information (MI) and Normalized Mutual Information (NMI). SSD, given in (2.1) is only suitable for monomodal images which differ only by gaussian noise and is susceptible if a few pixels have very large intensity differences.

$$SSD = \frac{1}{N} \sum_X (I_F(x) - I_M(T(x)))^2 \quad (2.1)$$

On the other hand, correlation coefficient and correlation ratios given in (2.2) and (2.3) assume some form of intensity relationship between the intensities of the fixed and moving image. This approach is also only suitable for monomodal images. In (2.3), $\bar{X}$ and $\bar{Y}$ are the random variables obtained from the intensity histograms of I_F and I_M .

$$CC = \frac{\sum_X (I_F(x) - \bar{I}_F) \cdot (I_M(T(x)) - \bar{I}_M)}{\sqrt{\sum_X (I_F(x) - \bar{I}_F)^2 \cdot \sum_X (I_M(T(x)) - \bar{I}_M)^2}} \quad (2.2)$$

$$\eta = 1 - \frac{Var[\bar{Y} - E(\bar{Y}|\bar{X})]}{Var(\bar{Y})} \quad (2.3)$$

Cost functions like Mutual Information (MI) given in (2.4) assume a probabilistic relationship between pixels, where $H(I_F)$ and $H(I_M)$ are the shanon entropy for the fixed and moving image and $H(I_F|I_M)$ is the conditional shanon entropy. Normalized Mutual Information (NMI) given in (2.5) solves the overlap problem with MI. MI is one of the popular methods for registering multi-modal images. Other intensity driven registration schemes have also included local frequency information [49].

$$MI = H(I_F) + H(I_M) - H(I_F|I_M) \quad (2.4)$$

$$NMI = \frac{H(I_F) + H(I_M)}{H(I_F|I_M)} \quad (2.5)$$

The performance of mutual information is very hard to predict in a number of different feature spaces and most algorithms would search and optimize for a local maxima [50]. Since registration itself is inherently a non-convex problem, cost functions like Mutual Information would aggravate the problem and increase non-convexity and non-linearity [51]. Mutual Information is also highly discretized [2] which was also shown in Fig. 1.3. On the other hand, fast and efficient optimization schemes rely on smooth and continuous functions.

Another drawback of MI is that it is dependent on the joint probability distribution of the intensities of two images. Since the probability density function is an approximation based on the number and size of bins or parzen window, this approximation would introduce errors into the calculations. Secondly, a meaningful joint probability distribution needs to have a large number of samples which would not be the case if a local region is being registered which would

have very few samples. Local nonrigid registration would therefore be constrained and mutual information would favor larger structures as can be seen in Fig. 2.1.

Too little variation in structure and underlying anatomy would make noise more dominant in an intensity based method. Since cost functions like MI completely disassociate the gray values of the image from their spatial location, there are many instances where registration could simply be driven by noise as very few structures are present in the image. Similarly, the loss of spatial information would also constrain the performance of MI for cases where symmetry is present. Brain MRIs and CT scans are frequently misregistered by MI algorithms because of the symmetric nature of the brain images. Symmetry in images could also fail optimization methods like stochastic gradient descent.

A lot of research has been done on intensity based methods and the solution to the nonrigid registration problem has only drifted away from simplification [46]. The introduction of an increasing number of parameters to improve and constrain intensity-based methods has made the solutions complex and it is difficult to think that research in intensity-based methods would now lead to a fast, efficient and accurate algorithm.

2.1.2 Feature-based Registration

Feature-based methods insert a feature extracting step before applying a similarity criterion to match features. Points, corners, land-marks, lines, curves and surfaces are generally considered good geometrical features for feature-based registration. Feature-based methods target the underlying anatomy and most of the times exclude plain or smooth areas within an image but seek to isolate salient and easily identifiable regions which are then used as land marks for matching purposes. When considering medical images, it is desirable that such features also represent distinct anatomical and functional landmarks. Linear features like ridges and crest lines have been used previously for nonrigid registration, [53], [54], [55], [56]. Features like crest lines can be intuitively appealing in some cases as they represent the sulci and gyri in the brain. A crest line atlas of the brain was constructed using these features in [57]. Surfaces could be similarly used as landmarks. 3D geometric features could be used to extract surfaces as landmarks, such as the surface of the left ventricle [58]. Brain surfaces were extracted in pre-operative brain MRI and CT scans using level set methods and then registration was done to track nonrigid deformations of the brain during surgery [59], [60].

The pre-processing required for feature extraction reduces the dimensionality of the problem as intensity information can now be discarded for calculating the transformation which helps in improving the computational speed. Feature extraction itself is far from reliable and anatomical and functional landmarks are difficult to define mathematically [61], [62]. In addition to this, finding proper similarity measures for feature matching and correspondence is also far from

FIGURE 2.1: Local registration is a problem with most nonrigid mutual information driven registration schemes. Results from Elastix Software [52] are shown. The top image shows the fixed image, the middle image is the registered image, and the bottom image is the moving image. Registration results show that small and local regions are not registered properly.

ideal [13], [63] and registration could be severely affected if incorrect correspondences are used as initial conditions.

Points are the building blocks for all other features. Good point features should be sparse, salient and well localized. Point registration schemes that do not focus on finding point correspondences first require dense point clouds without additional information to increase the statistical power of the spatial distribution of points [64]. On the other hand, if finding point correspondence is considered an important first step then high dimensional discriminative point descriptors are helpful. Steerable filters and Scale Invariant Feature Transform (SIFT) [65] have been used previously to construct high dimensional scale and rotation invariant descriptors. Moments have been widely used as well to construct invariant point descriptors [66].

Cognitive science suggests that edges and points are more robust and useful in identifying objects. Edge features are used by humans for visual recognition [5] and have been shown to be better at identifying objects than color or intensity [6]. Using points as features also helps in reducing memory and computational load which is a great advantage if 3D images are being registered. The presence of noise means that point features and point correspondence cannot result in exact matches. The presence of outliers is another problem faced by point registration scheme. The mismatch in the number of points would lead to some points being rejected in the matching stage as point correspondences need to be one-to-one. Lastly, the resulting geometric transformations would need to be responsive to highly elastic deformations and should avoid folding.

Point set registration is a two fold problem

- Find point correspondences between two point sets
- Estimate the transformation aligning the two sets of points

Existing point registration methods can be divided into three distinct groups.

- **Transformation Only:** Moments [67], Hough transform [68], Hausdorff distance [69], geometric hashing [70] and tree searches are used to find spatial transformations without explicitly finding point correspondences. These methods work well for rigid registration but are difficult to extend to nonrigid registration because of the complexity of the problem.
- **Correspondence Only:** Dense feature-based methods look for higher level structures by combining points, lines and curves. Modal matching [71] is another method which uses a mass and stiffness matrix built up from the gaussian of the distances between point features in either set. These methods are not robust to outliers. Graph matching [72] has also been used to find correspondences by constraining the search using spatial relationships.

- **Correspondence and Transformation:** Iterative Closest Point [13] and its variants [73], [74], [75], [76], [77], [78] are popular methods for solving point set registration problems. The problems associated with ICP are that it is sensitive to the initial conditions and is frequently trapped in local minima, especially for nonlinear image registration problems. Secondly, since the ICP is based on finding point correspondences, a threshold needs to be determined heuristically and errors can occur when describing what constitutes corresponding points.

To overcome ICP limitations, many probabilistic methods were developed. Methods used soft assignment to established point correspondences based on some probabilistic parameter [79], [80], [81], [82]. Although soft assignment methods provide an improvement for outliers over the general ICP but the solutions are still far from ideal.

In a similar manner, point set registration has been modeled as a maximum likelihood (ML) estimation problem where one point set is treated as Gaussian Mixture Model (GMM) centroids while the other is considered as data points. [83], [84], [85], [86], [87], [87], [88]. The Expectation-Maximization (EM) algorithm used for optimization of the likelihood functions consists of E-step to compute the probabilities and the M-step to update the transformation. Generally two probability distributions are constructed based on a gaussian mixture model (GMM) and a dissimilarity between the two is minimized. The drawback for such schemes is that outliers are not explicitly modeled and proper initialization is required.

Chamfer distance is the most commonly used similarity metric. Chamfer distance performance degrades with cluttered background and noise and is generally used because most algorithms search for approximate feature matches which does not require a need for calculating accurate euclidean distances. Hausdorff distance is a good discriminative similarity metric but does not treat point matching as a bipartite point matching problem. Partial Hausdorff distance [69] has been used to provide robustness against outliers. A further improvement was achieved by locally constraining the Hausdorff distance in [89] by introducing Weighted Hausdorff distance.

2.2 Shape Descriptors for Shape Recognition, Retrieval and Classification

Shape-based descriptor can be divided into two categories [90]; one is region-based and the other is contour-based. Contour-based descriptors generally focus on properties related to shape of the contours where as region-based descriptor gather intensity related features which help in describing the properties of the surrounding neighborhood. A good survey of local shape descriptors is given in [91] [92].

2.2.1 Region-based Feature Descriptor

Region-based or patch-based descriptors focus on extracting descriptive information from the surrounding neighborhood. Interesting points are extracted and features describing intensity and texture of the neighborhood patch are used as important discriminative features. SIFT [65] and geometric blur [93] have gained a lot of popularity in this class of feature descriptors. The problem with such patch-based descriptors is that the window size chosen is generally fixed, and the richness of the descriptors varies from region to region. Some patches might provide rich descriptive features where as other areas would have sparsely distributed informative features. Feature-based descriptors also suffer from ambiguity as features are locally defined and a global model or measure is therefore incorporated for feature matching. Euclidean norm in coordinate space is one important measure which is used to determine the accuracy of matching. Coarse to fine optimization is a better method for finding the optimal corresponding points based on gaussian scale space. A new technique described in [94] uses the *cosaliency* measure which determines feature correspondences at two levels. At the first level, feature vector correspondences are determined and at the second level the coherence of the matched points are determined spatially. Coherence in [94] is defined by the fitness of matched points relative to each other. For example, a square must still look like a square even after their point correspondences are matched and the spatial coherence should not be disturbed. Problems would still arise in such symmetric shapes as there would be more than one possibility for finding correspondences. Many intensity-based approaches build classifiers for shape recognition without finding explicit region or point correspondences [95] but these techniques would not be able to handle rotation, scale and translation variations.

Wavelets [96] and Gabor [97] vectors can extract features at many different scales and resolutions which makes them invariant to scale and resolution, unlike invariant moments where the size of the neighborhood is generally fixed. Gabor filters are sometimes used for multi-modal image registration which captures texture. Gabor function is not rotation invariant which is a problem for application such as image registration.

2.2.2 Contour-based Feature Descriptors

Contour-based descriptors generally extract contour properties like circularity, eccentricity, convexity, euler number, major axis orientation, concavity tree, principal axis ratio, circular variance, and elliptic variance [98]. They also focus on finding distinct and unique points on the contour as reference for matching. A popular category of contour-based shape descriptors uses shape signatures for finding point correspondences. Shape signatures take silhouette contour and represent it in 1D. Shape signatures generally have high matching and computational cost and are sensitive to noise [90]. The most popular signature-based shape descriptors are fourier

descriptors. Fourier descriptors [99] have been cited as being the best performing feature descriptors for silhouettes, where closed contours are used. But these descriptors would be hard to construct in noisy cluttered environment. Lower order harmonics are generally used to tackle nonrigid deformation but this prevents fourier descriptors from describing objects with sharp contour transitions. Shapes would not be clearly delineated and at most times multiple shapes would be existing together hence such an approach would fail in real world applications.

Another popular technique for shape matching uses point-based descriptors. Finding corresponding points using shape context has been previously described in [7], [14]. It assigns a feature descriptor for each interest point in an image which is then used to find a corresponding point in another image. The algorithm looks for descriptors that provide local as well as global shape context for each point. Some shape metrics like inner distance [100] might give good results when it is easier to extract contours but would not be robust to noise and outliers once contour detection becomes difficult and significant preprocessing would be required to extract a reliable object contour. The algorithm requires the application of a grid around a point which is then used to describe the neighborhood boundary points relative to it. Such algorithms are not robust as the grid used is rigid and slight deformations would push important descriptive contour points out of the grid. Similarly, other algorithms like [101] would be too dependent on the accuracy of contour detection for linking contours and would not give good results for real world images. Such algorithms would converge to local minimas and they would be less effective. An interesting algorithm was suggested in [102] where some regions of the contour were considered deformable. This method is intuitive and suggests that certain shapes would still remain identifiable to the human visual system even if certain regions of the shapes are deformed whereas they would be rendered unrecognizable if some other specific regions is deformed. This is also a training based approach and would not be easily generalizable.

Some algorithms use different measure for finding global and local shape descriptors. Different global and local features were used in [103]. Its local features are defined using object contours by gathering all the points in a circular window and then finding the angle that each point makes with the tangent angle of the point. The tangent angle in this case is defined as the tangent made with the contour. This is a very weak definition as the tangent might change due to deformation or due to missing contours which would render the descriptor useless and inaccurate. Its global feature is described by using the shape context technique [14]. Shape matching using shape contexts has shown good results for applications like character recognition but it lacks the robustness for nonlinear deformation.

Most of the work related to contour-based shape descriptors focuses on describing the shapes of closed contour objects normally referred to as silhouettes [104]. The boundary of such objects can be easily marked and most techniques which may work for well defined silhouettes may not work for real images where the boundary points are noisy and occluded. Extracting silhouettes

FIGURE 2.2: (a) describes the contour of a camel, (b) describes the contour with significant portions eliminated. Note that the elimination of significant contour points still makes the shape of the camel identifiable

in real images is a difficult process. Some algorithms have suggested techniques for handling missing contour information [105], suggesting that shapes are recognizable even when most of the contour points are missing as shown in Fig. 2.2. Although the descriptors derived are robust to occlusions, outliers and deformations, their performance is still far from ideal.

2.2.3 Shape Matching using Similarity Measures

Many similarity measures are available for finding matching contours or features [106]. Chamfer distance [107] has been employed as a similarity measure in many studies. Hausdorff distance [106] has also been a popular choice because of its discriminative power for shape recognition but the hausdorff distance may not be effective in noisy environments with many outliers. Partial Hausdorff distance considerably improves performance but its discriminative power for many different shapes remains questionable. Euclidean distance or the L_2 norm has been used in most studies to find similarity between features, but problems arise as all features may not have the same discriminatory power. Some shapes may be more easily identifiable using certain features, whereas other shapes would be dependent on some other specific feature in the feature vector. Learning schemes have been devised for the distant measure itself which has been trained on specific shapes [108]. The problem with such schemes is that they are not generic and sufficient training is required for accurate prediction. Such schemes may not work in natural settings where a large number of different shapes could be observed. Another method for learning schemes is by assigning weights to each feature after training. These schemes could be trained to be generic or specific for searching certain shapes. Some techniques for finding point correspondence [109] jointly optimize both alignment and shape matching, but these techniques suffer as they use distance similarity in the coordinate space to align the points and finding correspondence. Rich shape descriptors are not used, which would lead to error as points on dissimilar shapes might be matched to each other. Distance alone is not a good criterion for differentiating shapes and features describing the neighborhood intensity should also augment the closeness in coordinate space.

The task must therefore be to construct a scale, translation and rotation invariant shape descriptor from a combination of contour-based and region-based features, which would exhibit robustness against noisy contours and deformation and should be computationally less expensive. Statistical descriptors or moments are more suitable descriptors which can be robust to outliers and noisy contours.

2.3 Conclusion

In this literature review we provided a general framework for image registration and a brief overview of its different applications. An extensive survey on feature-based and intensity-based registration was provided. Shape descriptors for object classification were explored which provide a good criterion for correspondence. Extensive research related to nonrigid point-based registration has still not achieved desirable results. Correspondences are not well defined and accuracy of registration is greatly affected by outliers, occlusions and noise. The challenge is to come up with simple and robust algorithms which could function with good accuracy and speed even without a regularization framework.

Chapter 3

New Shape Descriptor for Object Classification

Describing shapes of objects in images is an important task in applications such as object recognition and classification. We present a novel shape descriptor based on principal curvatures and their directions. High curvature boundary points are first extracted in a multi-scale environment. The shape descriptor is extracted by taking the k nearest neighborhood boundary points around an interest point and calculating the central moments of the distribution of different properties of these neighborhood points. The choice of scale and the number of nearest neighbors provides a global as well as a local description of the neighborhood. Once point correspondences between images are found, a simple quaternion fitting algorithm [10] is used to align the boundary points of two different shapes. Euclidean distance is then computed to measure the similarity between matching points after alignment, both in the coordinate space and in the feature space for object recognition and classification. Compared to previous techniques for shape matching and classification, experiments show that the proposed algorithm is more robust to deformation and invariant to rotation, scale and translation. Table 3 lists some frequently used symbols and their description.

The contribution of our work on point-based shape descriptors is two fold:

- It introduces a new shape descriptor based on moment invariants extracted from principal curvatures and their directions. This provides a more compact description of shape which is not dependent on the missing boundary points or outliers and can handle significant deformation. It provides a coarse to fine multi-scale optimization which makes the feature descriptor robust to local deformations. Our algorithm is fast and can be further speeded up by implementing faster implementation of the nearest neighbor algorithm.

TABLE 3.1: frequently used symbols and their description

Symbol	Description
σ	standard deviation or scale of gaussian
(x, y)	coordinates of a point in image I
X	set of coordinates of all the points in image I
$(\cdot, \cdot)$	for all the coordinate points $(x, y) \in X$
P_j^σ	represents the coordinates (x, y) for j^{th} interest point at scale σ
$P_j^{\sigma_1, \sigma_2, i}$	represents the i^{th} nearest neighbor of point $P_j^{\sigma_1} \in S(\sigma_1)$ in set $S(\sigma_2)$
$k(\sigma)$	the number of nearest neighbors at scale σ
$\mathbf{P}_j^{\sigma_1, \sigma_2}$	represents $k(\sigma_2)$ nearest neighborhood points for point $P_j^{\sigma_1} \in S(\sigma_1)$ in set $S(\sigma_2)$
$\vec{p}_1(\sigma; x, y)$ & $\vec{p}_2(\sigma; x, y)$	maximum and minimum principal directions
$\mathbf{k}_1(\sigma; x, y)$ & $\mathbf{k}_2(\sigma; x, y)$	maximum and minimum principal curvatures

- It uniquely identifies symmetric (reflection) boundary points based on the feature descriptor itself without applying a coordinate space distance measure or any other global model to restrict the search for corresponding points. Other algorithms like fourier descriptors [110], [111] or shape context [14] are either unable to uniquely identify symmetric boundary points or are not robust to rotational variance.

3.1 Proposed Shape Descriptor

We construct our feature-based directional shape descriptor in two steps. The first step involves the extraction of interest points based on principal curvatures followed by a k nearest neighbor formulation where each interest point is represented by its k nearest neighbor points. The k nearest neighbors are extracted in a multi-scale environment to provide a global as well as local descriptor. The second step involves extracting rotation, scale and translation invariant feature descriptors and taking low order moments of those features.

3.1.1 Extracting High Curvature Interest Points

Shape of an object can be captured by a finite set of points. These points are either important landmark points or unique regions that can be easily identifiable on an object. The human visual system simplifies object recognition by searching for high curvature points [112]. These points tend to stand out in an image and mostly represent object contours. We also use a scale-space representation for extracting features as real world objects and shapes could be identifiable at many different scales [113]. The scale-space representation also provides us with a global

context which minimizes the possibility of finding local minimas. Two objects are shown in Fig. 3.1 from the Sclaroff tool database [114] which are present at different scales.

FIGURE 3.1: Similar objects are shown having different scales

A simple gaussian kernel $G(\sigma; x, y)$ given in (3.1) with standard deviation σ , is used to generate a scale-space representation $I(\sigma; \cdot, \cdot)$ of image $I(\cdot, \cdot)$ as shown:

$$G(\sigma; x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.1)$$

$$I(\sigma; \cdot, \cdot) = G(\sigma; \cdot, \cdot) * I(\cdot, \cdot) \quad (3.2)$$

We extract high curvature interest points by computing the principal curvatures and principal directions of the Hessian Matrix:

$$H(\sigma; x, y) = \begin{bmatrix} I_{xx}(\sigma; x, y) & I_{xy}(\sigma; x, y) \\ I_{yx}(\sigma; x, y) & I_{yy}(\sigma; x, y) \end{bmatrix} \quad (3.3)$$

where $I_{xx}(\sigma; x, y)$ represents the double derivative in the x direction, $I_{yy}(\sigma; x, y)$ represents the double derivative in the y direction, and $I_{xy}(\sigma; x, y)$ and $I_{yx}(\sigma; x, y)$ represent the derivative taken in the x and then y direction at scale σ . The eigenvectors and eigenvalues of the Hessian Matrix provides us with principal directions, $\vec{p}_1(\sigma; x, y)$ and $\vec{p}_2(\sigma; x, y)$ and their corresponding principal curvatures $\mathbf{k}_1(\sigma; x, y)$ and $\mathbf{k}_2(\sigma; x, y)$. The first eigenvector $\vec{p}_1(\sigma; x, y)$ gives the direction of greatest curvature where as its corresponding eigenvalue $\mathbf{k}_1(\sigma; x, y)$ representing the amount of curvature, the second eigenvector $\vec{p}_2(\sigma; x, y)$ represents the direction of least curvatures and its eigenvalue $\mathbf{k}_2(\sigma; x, y)$ representing the amount of curvature. The principal curvature vectors and their intensities are useful in differentiating between different types of interest points. If $\mathbf{k}_1(\sigma; x, y) \approx \mathbf{k}_2(\sigma; x, y)$, then the intensity distribution around the points is spherical, where as if $\mathbf{k}_1(\sigma; x, y) \gg \mathbf{k}_2(\sigma; x, y)$ then the intensity distribution around the point

FIGURE 3.2: Principal curvature weighted principal directions.

FIGURE 3.3: Principal curvature points extracted at three different scales (Left: $\sigma = 1$, Middle: $\sigma = 2$, Right: $\sigma = 4$) to represent the boundary of the hammer object taken from the Sclaroff data set.

would most likely represent a ridge with the principal vector $\vec{p}_1(\sigma; x, y)$ pointing in the direction of the ridge. Similarly, principal directions and curvatures can be used to differentiate saddle points and corners.

The principal curvature weighted principal directions $\vec{p}_1(\sigma; \cdot, \cdot)$ are shown in Fig. 3.2. The image $I(\sigma; \cdot, \cdot)$ was first thresholded by removing 95% of the unwanted low curvature points. A threshold was determined by $T_{thres} = \mu(\mathbf{k1}(\sigma; \cdot, \cdot)) + 2\sigma(\mathbf{k1}(\sigma; \cdot, \cdot))$, where μ and σ represent mean and standard deviation. We obtain a set of n points $S(\sigma) = \{P_1^\sigma, P_2^\sigma, \dots, P_n^\sigma\}$ for image I with scale σ , where P_j^σ represents the j^{th} interest point in a 2D coordinate space.

Three sets of interest points are extracted for each image at scale $\sigma = 1, 2$ and 4 . The value of σ is dependent on the size of the images and to generate a regular scale-space pyramid, σ should exponentially increase [113]. We treat the three sets of interest points, $S(1)$, $S(2)$, and $S(4)$ separately and extract features using the scale at which the interest point P_j^σ was extracted. $S(1)$, $S(2)$, and $S(4)$ are plotted for an object taken from Sclaroff tool database in Fig. 3.3.

FIGURE 3.4: Interest points shown describing the principal curvatures and the 8 nearest neighbor points distributed around the interest point. Principal curvature directions are represented by the vector at $P_j^{\sigma_1, \sigma_2, i}$

For each point $P_j^{\sigma_1} \in S(\sigma_1)$ we search for $k(\sigma_2)$ nearest neighborhood points in $S(\sigma_2)$ given by $\mathbf{P}_j^{\sigma_1, \sigma_2}$:

$$\mathbf{P}_j^{\sigma_1, \sigma_2} = \{P_j^{\sigma_1, \sigma_2, 1}, P_j^{\sigma_1, \sigma_2, 2}, \dots, P_j^{\sigma_1, \sigma_2, k(\sigma_2)}\} \quad (3.4)$$

where $P_j^{\sigma_1, \sigma_2, i}$ represents the i^{th} nearest neighbor of point $P_j^{\sigma_1} \in S(\sigma_1)$ in set $S(\sigma_2)$. Replacing each interest point $P_j^{\sigma_1} \in S(\sigma_1)$ by its neighborhood point set $\mathbf{P}_j^{\sigma_1, \sigma_2}$, we obtain $\mathbf{S}(\sigma_1, \sigma_2)$:

$$\mathbf{S}(\sigma_1, \sigma_2) = \{\mathbf{P}_1^{\sigma_1, \sigma_2}, \mathbf{P}_2^{\sigma_1, \sigma_2}, \dots, \mathbf{P}_n^{\sigma_1, \sigma_2}\} \quad (3.5)$$

Feature vectors are only constructed for the finer scale interest points present in $S(1)$ where as the higher scale points provide a global context to these finer scale points. Different values for $k(\sigma)$ are used at each scale to derive a multi-scale nearest neighbor point set $\bar{\mathbf{S}}$ for each point $P_j^1 \in S(1)$ as shown in (3.6):

$$\bar{\mathbf{S}} = \begin{bmatrix} \mathbf{S}(1, 1) & \mathbf{S}(1, 2) & \mathbf{S}(1, 4) \end{bmatrix} \quad (3.6)$$

3.1.2 Feature Extraction

After calculating $\bar{\mathbf{S}}$ and finding the nearest neighbors associated with each point P_j^σ we focus on extracting rotation, translation and scale invariant features that will also be robust to deformation. Fig. 3.4 gives a description of how $k(\sigma)$ nearest neighbor interest points would be located around a point $P_j^\sigma \in S(\sigma)$.

Various studies have been done to extract moment invariant feature descriptors for image matching and classification [115], [116]. A number of geometric invariant primitives are presented in [116] which are invariant to rotation and translation. The 2D geometric primitives include distance D , area A and dot product d (see Appendix A). These geometric primitives can be made scale invariant by normalizing them to one.

We extend this concept of geometric primitives and search for image properties which could be combined with these primitives to give a more discriminative feature descriptor. For every point $P_j^{\sigma_1}$ we would have a multi-scale nearest neighbor point set $\mathbf{P}_j^{\sigma_1, \sigma_2}$ shown in Fig. 3.4. The first geometric primitive, distance D , would be easy to extract. This distance is calculated between each of the multi-scale nearest neighborhood points in $\mathbf{P}_j^{\sigma_1, \sigma_2}$ and reference point $P_j^{\sigma_1}$. The resulting distance vector $\mathbf{D}_j^{\sigma_1, \sigma_2}$, shown in (3.7), is obtained for each multi-scale nearest neighborhood point set $\mathbf{P}_j^{\sigma_1, \sigma_2} \in \bar{\mathbf{S}}$.

$$\mathbf{D}_j^{\sigma_1, \sigma_2} = \begin{bmatrix} \|P_j^{\sigma_1} - P_j^{\sigma_1, \sigma_2, 1}\| \\ \|P_j^{\sigma_1} - P_j^{\sigma_1, \sigma_2, 2}\| \\ \vdots \\ \|P_j^{\sigma_1} - P_j^{\sigma_1, \sigma_2, k(\sigma_2)}\| \end{bmatrix}^T \quad (3.7)$$

The choice of a reference axis is an important step for feature extraction. A suitable reference axis should be able to adjust with rotation to make the feature descriptor invariant. Many shape descriptors use the coordinate space axis as the reference axis which makes the feature descriptor ineffective for rotational variance. The contour tangent is a popular choice for reference axis [14]. This approach makes the shape descriptor invariant to rotation but is not robust as a slight change in the tangent vector would produce drastic changes in the calculated descriptors. Contour points could be deformed or occluded which would result in incorrect descriptors. Another approach uses the principal component of the contours as the reference axis but would also suffer when there are missing or occluded points.

Since the choice of the feature descriptor plays an extremely important role in making our feature vector robust, we use scale-space to refine our choice. Our candidate for the choice of a reference axis is the principal curvature vector $\vec{p}_1(\sigma; x, y)$ which is an inherent image property and would not be dependent on the contour of the object. We use a coarse scale to extract the principal

vector $\vec{p}_1(\bar{\sigma}; \{P_j^\sigma\})$ at the interest point P_j^σ , where $\bar{\sigma}$ represents the coarsest scale. The principal vector would remain unchanged even if points are occluded or missing. In the case where neighborhood points are chosen as the interest point, the principal direction would still be similar to the original interest point in the coarse scale.

We use the reference axis $\vec{p}_1(\bar{\sigma}; \{P_j^\sigma\})$ at each point P_j^σ to calculate the dot product which, as discussed earlier, is a rotation, translation and scale invariant primitive. The dot products are calculated between the reference axis, $\vec{p}_1(\bar{\sigma}; \{P_j^{\sigma_1}\})$, at point $P_j^{\sigma_1}$ and the principal curvature vector of each nearest neighbor interest point $P_j^{\sigma_1, \sigma_2, i} \in \mathbf{P}_j^{\sigma_1, \sigma_2}$ to obtain $\mathbf{d}_j^{\sigma_1, \sigma_2}$ as shown in (3.9). The dot product $d(m, n, \sigma_1, \sigma_2)$ is taken using unit vectors which is equivalent to the cosine of the inner angle θ subtended by the two principal curvature vectors $\vec{p}_1(\bar{\sigma}; \{P_j^{\sigma_1}\})$ and $\vec{p}_1(\sigma_2; \{P_j^{\sigma_1, \sigma_2, i}\})$ as shown in (3.8). The cosine of an angle θ fails to differentiate between vectors that subtend a positive θ from vectors that subtend a negative θ . Hence any distribution of neighborhood principal vectors $\vec{p}_1(\sigma; \{\mathbf{P}_j^{\sigma_1, \sigma_2}\})$ if reflected along the direction of $\vec{p}_1(\bar{\sigma}; \{P_j^{\sigma_1}\})$ would produce the same vector $\mathbf{d}_j^{\sigma_1, \sigma_2}$.

$$d(j, i, \sigma_1, \sigma_2) = \vec{p}_1(\sigma_1; \{P_j^{\sigma_1}\}) \bullet \vec{p}_1(\sigma_2; \{P_j^{\sigma_1, \sigma_2, i}\}) \quad (3.8)$$

$$\mathbf{d}_j^{\sigma_1 \sigma_2} = \begin{bmatrix} d(j, 1, \sigma_1, \sigma_2) \\ d(j, 2, \sigma_1, \sigma_2) \\ \vdots \\ d(j, k(\sigma), \sigma_1, \sigma_2) \end{bmatrix}^T \quad (3.9)$$

To further enhance the discriminatory power of our feature descriptor we then extract the cross product using the curvature weighted principal directions $\mathbf{k1}(\sigma; \{P_j^\sigma\}) * \vec{p}_1(\sigma; \{P_j^\sigma\})$. Cross product, as described previously, is a rotation, translation and scale invariant primitive. The cross product $\Lambda(j, i, \sigma_1, \sigma_2)$ is given by:

$$\Lambda(j, i, \sigma_1, \sigma_2) = \left(\mathbf{k1}(\bar{\sigma}; \{P_j^{\sigma_1}\}) * \vec{p}_1(\bar{\sigma}; \{P_j^{\sigma_1}\}) \right) \times \left(\mathbf{k1}(\sigma_2; \{P_j^{\sigma_1, \sigma_2, i}\}) * \vec{p}_1(\sigma_2; \{P_j^{\sigma_1, \sigma_2, i}\}) \right) \quad (3.10)$$

The cross product in (3.10) is taken between the curvature weighted reference axis $\mathbf{k1}(\bar{\sigma}; \{P_j^{\sigma_1}\}) * \vec{p}_1(\bar{\sigma}; \{P_j^{\sigma_1}\})$ at interest point $P_j^{\sigma_1}$, and its i^{th} nearest neighbor curvature weighted principal direction $\mathbf{k1}(\sigma_2; \{P_j^{\sigma_1, \sigma_2, i}\}) * \vec{p}_1(\sigma_2; \{P_j^{\sigma_1, \sigma_2, i}\})$. Cross product $\Lambda(j, i, \sigma_1, \sigma_2)$ is capable of differentiating neighborhood principal directions $\vec{p}_1(\sigma_1; \{P_j^{\sigma_1, \sigma_2, i}\})$, if reflected along the direction of $\vec{p}_1(\bar{\sigma}; \{P_j^{\sigma_1}\})$. We have deviated from the definition of $A(p, q, r)$ (see Appendix B)

by not taking the minkowski distance ($L2$ norm) $\|\cdot\|_2$ so that $\Lambda(j, i, \sigma_1, \sigma_2)$ differentiates between symmetric principal curvature vector distributions by giving negative or positive outputs. This is intuitively appealing as $\vec{p}_1(\sigma; \{P_{j_1}^\sigma\}) \times \vec{p}_1(\sigma; \{P_{j_2}^\sigma\}) \neq \vec{p}_1(\sigma; \{P_{j_2}^\sigma\}) \times \vec{p}_1(\sigma; \{P_{j_1}^\sigma\})$. This removes the need for introducing a global perimeter to maintain spatial coherence between matched points especially if interest point distribution has reflective symmetry. As can be seen in Fig. 3.5, our feature descriptor is capable of discriminating points lying on the right side of the hammer boundary from the left. Although completely symmetric objects like a circle, or reflective and rotational symmetries of higher order, would require a global perimeter or model, but most natural objects are not perfectly symmetric and at best represent lower order symmetries. Cross product $\Lambda(j, i, \sigma_1, \sigma_2)$ is calculated for all points $P_j^{\sigma_1, \sigma_2, i} \in \mathbf{P}_j^{\sigma_1, \sigma_2}$ to give $\Phi_j^{\sigma_1, \sigma_2}$ (3.11).

$$\Phi_j^{\sigma_1, \sigma_2} = \begin{bmatrix} \Lambda(j, 1, \sigma_1, \sigma_2) \\ \Lambda(j, 2, \sigma_1, \sigma_2) \\ \dots \\ \Lambda(j, k(\sigma_2), \sigma_1, \sigma_2) \end{bmatrix} \quad (3.11)$$

Another important feature extracted are the eigenvalues, $\mathbf{k1}(\sigma; \cdot, \cdot)$ and $\mathbf{k2}(\sigma; \cdot, \cdot)$, of the Hessian matrix given in (3.3). These principal curvatures provide a measure for the distribution of the underlying intensity values. They would help to differentiate ridges, corners and blobs. The principal curvatures, $\mathbf{k1}(\sigma; \cdot, \cdot)$ and $\mathbf{k2}(\sigma; \cdot, \cdot)$, are extracted for all the $k(\sigma_2)$ nearest neighborhood points $P_j^{\sigma_1, \sigma_2, i} \in \mathbf{P}_j^{\sigma_1, \sigma_2}$ to give $\Omega \mathbf{1}_j^{\sigma_1, \sigma_2}$ and $\Omega \mathbf{2}_j^{\sigma_1, \sigma_2}$ as shown:

$$\Omega \mathbf{1}_j^{\sigma_1, \sigma_2} = \begin{bmatrix} \mathbf{k1}(\sigma_2; P_j^{\sigma_1, \sigma_2, 1}) \\ \mathbf{k1}(\sigma_2; P_j^{\sigma_1, \sigma_2, 2}) \\ \vdots \\ \mathbf{k1}(\sigma_2; P_j^{\sigma_1, \sigma_2, k(\sigma_2)}) \end{bmatrix}^T \quad (3.12)$$

$$\Omega \mathbf{2}_j^{\sigma_1, \sigma_2} = \begin{bmatrix} \mathbf{k2}(\sigma_2; P_j^{\sigma_1, \sigma_2, 1}) \\ \mathbf{k2}(\sigma_2; P_j^{\sigma_1, \sigma_2, 2}) \\ \vdots \\ \mathbf{k2}(\sigma_2; P_j^{\sigma_1, \sigma_2, k(\sigma_2)}) \end{bmatrix}^T \quad (3.13)$$

At the end of our feature extraction step we would have a unique set of feature vectors for each point $P_j^1 \in \bar{\mathbf{S}}$. These feature vectors are $\mathbf{D}_j^{1,1}, \mathbf{D}_j^{1,2}, \mathbf{D}_j^{1,4}, \mathbf{d}_j^{1,1}, \mathbf{d}_j^{1,2}, \mathbf{d}_j^{1,4}, \Phi_j^{1,1}, \Phi_j^{1,2}, \Phi_j^{1,4}, \Omega \mathbf{1}_j^{1,1}, \Omega \mathbf{1}_j^{1,2}, \Omega \mathbf{1}_j^{1,4}, \Omega \mathbf{2}_j^{1,1}, \Omega \mathbf{2}_j^{1,2}$, and $\Omega \mathbf{2}_j^{1,4}$. We then use moment descriptors to describe these feature vectors to construct a more compact feature vector.

3.1.3 Moment-based Shape Descriptors

Most shape descriptors rely on the coordinate space for extracting descriptors. Shape context [14] uses radial bins to collect the boundary points using a reference axis. Fourier descriptors [99] [110] also use measures such as distance and angles which are strictly related to the location of the points in the coordinate space relative to some reference point. Moment based shape descriptors [116] have the advantage of characterizing both shape and intensity distribution around a point. We leverage this property of moment invariants and extract features. Higher order moments are generally more sensitive to noise and small deformations, we therefore restrict our analysis to low order moments. We use mean $\mu(X)$, variance $var(X)$, and the 3rd and 4th standardized moments about the mean, $\delta(X)$ and $\gamma(X)$, to characterize the feature distribution (see Appendix B). After extracting different nearest neighbor feature vectors for each interest point $P_j^1 \in \bar{\mathbf{S}}$, we then calculate the feature vector's μ , var , γ and δ values.

$$\overline{\mathbf{D}}_j^{\sigma_1, \sigma_2} = [\mu(\mathbf{D}_j^{\sigma_1, \sigma_2}), var(\mathbf{D}_j^{\sigma_1, \sigma_2}), \delta(\mathbf{D}_j^{\sigma_1, \sigma_2}), \gamma(\mathbf{D}_j^{\sigma_1, \sigma_2})] \quad (3.14)$$

$$\overline{\mathbf{d}}_j^{\sigma_1, \sigma_2} = [\mu(\mathbf{d}_j^{\sigma_1, \sigma_2}), var(\mathbf{d}_j^{\sigma_1, \sigma_2}), \delta(\mathbf{d}_j^{\sigma_1, \sigma_2}), \gamma(\mathbf{d}_j^{\sigma_1, \sigma_2})] \quad (3.15)$$

$$\overline{\mathbf{\Omega 1}}_j^{\sigma_1, \sigma_2} = [\mu(\mathbf{\Omega 1}_j^{\sigma_1, \sigma_2}), var(\mathbf{\Omega 1}_j^{\sigma_1, \sigma_2}), \delta(\mathbf{\Omega 1}_j^{\sigma_1, \sigma_2}), \gamma(\mathbf{\Omega 1}_j^{\sigma_1, \sigma_2})] \quad (3.16)$$

$$\overline{\mathbf{\Omega 2}}_j^{\sigma_1, \sigma_2} = [\mu(\mathbf{\Omega 2}_j^{\sigma_1, \sigma_2}), var(\mathbf{\Omega 2}_j^{\sigma_1, \sigma_2}), \delta(\mathbf{\Omega 2}_j^{\sigma_1, \sigma_2}), \gamma(\mathbf{\Omega 2}_j^{\sigma_1, \sigma_2})] \quad (3.17)$$

$$\overline{\mathbf{\Phi}}_j^{\sigma_1, \sigma_2} = [\mu(\mathbf{\Phi}_j^{\sigma_1, \sigma_2}), var(\mathbf{\Phi}_j^{\sigma_1, \sigma_2}), \delta(\mathbf{\Phi}_j^{\sigma_1, \sigma_2}), \gamma(\mathbf{\Phi}_j^{\sigma_1, \sigma_2})] \quad (3.18)$$

The μ , var , δ and γ values are collected and placed in a feature vector $\mathbf{F}_j(\sigma_1, \sigma_2)$ which calculates these four properties for different features extracted from the neighborhood points set $\mathbf{P}_j^{\sigma_1, \sigma_2}$ for point $P_j^{\sigma_1}$.

$$\mathbf{F}_j(\sigma_1, \sigma_2) = [\overline{\mathbf{D}}_j^{\sigma_1, \sigma_2}, \overline{\mathbf{d}}_j^{\sigma_1, \sigma_2}, \overline{\mathbf{\Omega 1}}_j^{\sigma_1, \sigma_2}, \overline{\mathbf{\Omega 2}}_j^{\sigma_1, \sigma_2}, \overline{\mathbf{\Phi}}_j^{\sigma_1, \sigma_2}] \quad (3.19)$$

To implement coarse to fine optimization and providing a global as well as a local descriptor for the shape relative to each point P_j^σ we use three different values for scale σ and $k(\sigma)$ nearest neighbors. The choice of $k(\sigma)$ at different scales was calculated experimentally and is shown in Fig. 4.1. The scales chosen as discussed earlier were $\sigma = 1, 2$, and 4 . This gives a 60 dimensional feature vector $\vec{\mathbf{F}}_j$ for each point P_j^1 .

$$\bar{\mathbf{F}}_j = [\mathbf{F}_j(1, 1), \mathbf{F}_j(1, 2), \mathbf{F}_j(1, 4)] \quad (3.20)$$

3.1.4 Feature Normalization

To get an accurate measure of similarity, the feature vector space needs to be normalized so that each feature contributes equally to the feature descriptor. We use a standard zero-mean and unit-variance (MV) normalization technique [117]. Most of the feature descriptors extracted represent a normal distribution hence this technique works well. Feature vector $\bar{\mathbf{F}}_j$ extracted for all the n interest points in image I is given below. Each column of $\bar{\mathbf{F}}$ represents a unique feature calculated for all the n interest points.

$$\bar{\mathbf{F}} = \begin{bmatrix} \bar{\mathbf{F}}_1 \\ \bar{\mathbf{F}}_2 \\ \vdots \\ \bar{\mathbf{F}}_n \end{bmatrix} \quad (3.21)$$

We normalize each column of $\bar{\mathbf{F}}$ using the MV normalization technique:

$$\bar{\mathbf{F}}^N = \begin{cases} \frac{\bar{\mathbf{F}}^c - \mu(\bar{\mathbf{F}}^c)}{\sigma(\bar{\mathbf{F}}^c)} & \forall c = 1, 2, \dots, 60 \end{cases} \quad (3.22)$$

where $\bar{\mathbf{F}}^c$ represents the c^{th} column of the feature vector $\bar{\mathbf{F}}$, and $\mu(\bar{\mathbf{F}}^c)$ and $\sigma(\bar{\mathbf{F}}^c)$ represent the mean and variance for the c^{th} column. Feature vector for interest point P_j^σ with all features normalized is given by $\bar{\mathbf{F}}_j^N$.

3.2 Object Classification

Feature descriptors are calculated for all the interest points in each image I . Different number of interest points n are obtained as each image varies from the other. We first look for point correspondences between images. The two feature sets shown in (3.23) contain feature descriptors calculated for all the interest points in the respective images. A simple nearest neighbor algorithm using euclidean norm gives us the closest matches in feature space between the two sets of feature descriptors, $\bar{\mathbf{F}}^N(I_1)$ and $\bar{\mathbf{F}}^N(I_2)$, as shown in (3.24) where j_1 and j_2 represent the index of the interest points extracted from two different images I_1 and I_2 . For each feature vector $\bar{\mathbf{F}}_{j_1}^N$ in I_1 , the closest feature vector $\bar{\mathbf{F}}_{j_2}^N$ is searched for in $\bar{\mathbf{F}}^N(I_2)$. The nearest neighbor search is performed iteratively to avoid “one to many” correspondences and to make sure that

each point is matched to the closest match in the other image and vice versa. The mismatch in the number of points n in different images would lead to many points without any matches. We will ignore all such points. We store the spatial coordinates of the matched points corresponding to the matched feature vectors in M , where coordinate $[x_{I_1}^r, y_{I_1}^r]$ represents the coordinates of the r^{th} point in image I_1 , and $[x_{I_2}^r, y_{I_2}^r]$ represents the coordinates of the corresponding matched point in I_2 .

$$\bar{\mathbf{F}}^N(I_1) \Leftrightarrow \bar{\mathbf{F}}^N(I_2) \quad (3.23)$$

$$j_2^* = \arg \min_{j_2} \|\bar{\mathbf{F}}_{j_1}^N(I_1) - \bar{\mathbf{F}}_{j_2}^N(I_2)\|_2 \quad (3.24)$$

$$M(r) = \{[x_{I_1}^r, y_{I_1}^r], [x_{I_2}^r, y_{I_2}^r]\} \quad (3.25)$$

(a)

(b)

FIGURE 3.5: Point correspondences between two hammer objects from the Kimia Tool Database, and two articulated scissor objects. Some points are mismatched due to the presence of noise and deformation but these mismatches are few in number.

The results for the point correspondences calculated can be seen in Fig. 3.5. A majority of the point correspondences are correct with few mismatches. Our shape descriptor is able to differentiate between points lying on symmetric boundaries as is the case in both the hammer

and scissor objects in Fig. 3.5. It is observed that a majority of mismatch points are the ones lying on low curvature boundaries.

3.2.1 Image Alignment using Matched Points

After finding point correspondence using the shape descriptors defined earlier we implement a standard method for object classification. Many shape classification problems [14], [110] use alignment as a first step towards classification after finding corresponding points. Alignment for our case is simple as many Iterative Closest Point (ICP) algorithms exist in closed form. We use one such algorithm based on quaternion rotation [10].

A quaternion based Iterative Closest Point (ICP) algorithm for registering two sets of points is proposed in [10]. Let $M(r)$ be the set of corresponding points to be aligned obtained from (3.25). Both sets $[x_{I_1}^r, y_{I_1}^r]$ and $[x_{I_2}^r, y_{I_2}^r]$ contain equal number of points after all the unmatched points are removed. The unit quaternion representing the best rotation is obtained by extracting the eigenvector of the 4×4 matrix $\mathbf{Q}_{MAT}$ given in (39) of Appendix C.

3.2.2 Outlier Removal and Iteration

To further refine the results of the ICP algorithm, we make the alignment into a two step process. After the initial alignment we eliminate 10% of the outliers based on the distance in coordinate space between matched points. These outliers are removed from the list of matched points and the alignment process is repeated.

Two hammer objects from the Kimia database are aligned in Fig. 3.6(a) using the Iterative Closest Point algorithm described above. The point correspondences shown in Fig. 3.5 were used as inputs to the quaternion matrix $\mathbf{Q}_{MAT}$. Our point correspondence and alignment algorithm works well as seen from the figure.

3.2.3 Cost Function

$$C_{i_1, i_2} = \sum_{j=1}^n [(\bar{x}_{i_1}^{j1} - x_{i_2}^{j2})^2 + (\bar{y}_{i_1}^{j1} - y_{i_2}^{j2})^2]^{1/2} \quad (3.26)$$

After aligning matched points $[x_{I_1}^r, y_{I_1}^r]$ with $[x_{I_2}^r, y_{I_2}^r]$, we get a new set of transformed points $[\bar{x}_{I_1}^r, \bar{y}_{I_1}^r]$. We define a simple cost function C_{i_1, i_2} based on the Euclidean distance and is given in (3.26). Observe that the cost function does not take scale variance into consideration although

FIGURE 3.6: Four different cases of aligning object based on the proposed shape descriptor using quaternion fitting are shown for the Sclaroff tool database. 3.6(a) and 3.6(b) show two different hammer contours before and after alignment. 3.6(c) and 3.6(d) show a hammer being matched to a double-ended wrench. 3.6(e) and 3.6(f) show before and after alignment contours for two different hammers having different scale. 3.6(g) and 3.6(h) show the alignment for two hammers where one has a different scale and is considerably deformed.

the point matching and alignment process is scale invariant. The coordinates, $M(r)$, are normalized for both sets of points $[\bar{x}_{I_1}^r, \bar{y}_{I_1}^r]$ and $[x_{I_2}^r, y_{I_2}^r]$ using mean-variance normalization before calculating C_{i_1, i_2} . The cost function given is therefore sufficient to handle scale variations and tolerates considerable deformation.

3.3 Conclusion

In this Chapter, we presented a theoretical framework for the construction of a new point-based shape descriptor. The shape descriptor is constructed using high curvature interest points. A multi-scale k nearest neighbor approach is used to collect neighborhood interest points and geometric moment invariants are then extracted from these points using principal curvatures and their directions. Corresponding points are then found using the $L2$ norm based on the feature vector. Once one-to-one correspondences are obtained, quaternion fitting is used to find the best transformation. The distance of points from the center is normalized to make the registration scale invariant. In Chapter 4 we present performance and comparison studies for this new shape descriptor.

Chapter 4

Performance Studies and Comparisons

We evaluate the proposed point-based shape descriptor given in Chapter 3 on the Sclaroff tool dataset [114], Kimia silhouettes data set [118], Articulated tool data set [119] and MNIST data set for handwritten digits [120]. These data sets have widely been used as benchmarks in many different studies and the proposed algorithm outperforms others in most of the comparative studies done using these databases.

4.1 Choice of Various Parameters for Optimizing Shape Matching

Experiments are performed to optimize the value of various parameters. The scale σ and the choice of the number of nearest neighbors k affect the accuracy of the point-based shape descriptor.

4.1.1 Choice of "k" for Nearest Neighbor Search for Feature Extraction

Our algorithm is dependent on the choice of the number of nearest neighborhood points k_1 , k_2 and k_3 , for scales σ_1 , σ_2 , and σ_3 . Although our experiment has shown that $k_2 = k_1 \times 2$ and $k_3 = k_2 \times 2$ provide the best results, for certain shapes which have higher details and local information, the choice of the number of nearest neighbor k_1 , k_2 and k_3 needs to be fine tuned. The "fighter plane" object from the Kimia database is one such object which has more fine scale detail. Hence the performance of the proposed algorithm increases drastically when k_1 is reduced further by a factor of 4 with respect to k_2 . Increasing the number of scales could improve the results for this particular image.

FIGURE 4.1: The values of k_1 , k_2 and k_3 were chosen empirically, where k_1 is the number of nearest neighbors for $\sigma = 4$, k_2 for $\sigma = 2$ and k_3 for $\sigma = 1$. In the experiment, ‘n’ refers to the total number of interest points extracted for each image. We use the Kimia tool database to run our experiments. The misclassification error rate was plotted and it can be seen that the proposed algorithm performed best for $k_1 = n/8$, $k_2 = n/16$ and $k_3 = n/32$.

4.1.2 Choice of Number of Scales σ

A number of objects present themselves at more scales than 3. The Cat object in the Kimia 99 database is one such image which has a lot more finer details. Although most 4 legged animals look the same but the image search is much improved if the finest resolution i.e. $\sigma = 0$ is also included. For large images, σ should be large. Since we have mostly experimented with small images whose resolution was less than 150×150 , the choice for $\sigma = 1, 2$ and 4 was sufficient. At $\sigma > 4$ the image would be blurred beyond recognition and its shape would hardly be distinguishable from other objects. $\sigma < 1$ is also not used as a number of noisy features are not filtered out and the computational cost increases due to a larger number of interest points. For a more generic approach we can link the choice of σ with the resolution of the image whose features are to be extracted.

4.2 Studies on Different Databases

4.2.1 Sclaroff Tool Database

We tested the proposed algorithm on the Sclaroff tool database [114] which contains 63 monochrome images of tools. Three different classes of tools; hammers, single-ended wrench and double-ended wrench are present. Each category contains 21 images each of size 128×120 . All the three categories are very similar and the differences in shape are minimal. Experiments were done by taking each image from the database and searching for the nearest matches in a leave one out fashion. The results from this database would help us evaluate the effectiveness of the

algorithm in differentiating similar objects, results are shown in Fig. 4.2. The best results are for the hammer and single-end wrench with 17 correct matches each out of 20 images. The double-ended wrench retrieval results include six incorrect matches out of 20, whereas there were only two incorrect matches in the top 10 nearest matches. Note that images for the double-ended wrench were comparatively noisier and the object was significantly more deformed. Table 4.1 shows the results for the first, second and third nearest matches for each query image.

TABLE 4.1: Retrieval accuracy of the first 3 nearest matches for the sclaroff tool database

Algorithm	Nearest Neighbor		
	first	second	third
hammer	20/20	20/20	20/20
single-ended wrench	20/20	19/20	19/20
double-ended wrench	19/20	19/20	18/20

4.2.2 Kimia-99 database

The Kimia-99 database [121] with nine categories and eleven shapes per category has been frequently cited by researchers studying shape recognition. The database provides varying number of different shapes with occlusions, articulation and deformations . We applied a leave one out strategy to find the nearest matches for each image in the database.

TABLE 4.2: Recognition accuracy of different algorithms compared with our proposed shape matching algorithm for Kimia-99 database

	Algorithm						
	KDE	SIHA	SFHM	KDE-ICA	IDSC+DP	Shape Context	Proposed
Top1	97	97	96	94	99	97	99
Top2	96	86	90	84	99	91	99
Top3	89	86	86	76	99	88	99
Top4	90	83	79	76	98	85	97
Top4	80	75	77	68	98	84	96

Table 4.2 shows the retrieval accuracy of top five nearest matches for 6 different algorithms. Kernel Density Estimation(KDE), Statistical Integration of Histogram Array (SIHA), Structural Feature Histogram Matrix (SFHM), Kernel Density Estimation with ICA (KDEICA), Inner Distance Shape Context with Dynamic Programming (IDSC+DP) [121] and Shape Context [14] were used for comparison purposes. The results for these algorithms were previously cited in [122]. Compared to these algorithms the accuracy of our shape retrieval algorithm is better, the proposed method correctly identifies all the top three nearest matches. The first few nearest matches were 100 percent accurate and the results were comparable to the one's given for IDSC+DP [121]. On the other hand, IDSC+DP [121] is dependent on closed object contours and is computationally more expensive and therefore has limited application beyond silhouette

(a)

(b)

(c)

FIGURE 4.2: The top twenty nearest matches are shown for three sample query images from the Sclaroff tool database; (a) hammer, (b) double-ended wrench, and (c) single-ended wrench. Observe that retrieval results include tool objects which were significantly deformed.

images. The results for our proposed algorithm are shown in Fig. 4.3. Fig. 4.4 shows some of the results for objects that are occluded or having missing boundary points. Kimia-99 database contains many deformed and occluded objects and the results show that the proposed algorithm performs better and is more robust compared to other algorithms.

FIGURE 4.3: First five retrieval results for the Kimia 99 database are given for each class. The images in the left most column are the query images and the five nearest matches are shown to their right.

FIGURE 4.4: Results for occluded objects of Kimia-99 database.

4.2.3 Kimia-25 Database

We also tested the proposed algorithm on the Kimia-25 database [118] which contains 25 images and 6 object categories. Table 4.3 summarizes the results of matching images of this database. The results are comparable to IDSC+DP [121].

TABLE 4.3: Retrieval accuracy for algorithms experimented on the Kimia-25 database. Results for the first three nearest matches retrieved are shown.

	Nearest Neighbor		
	first	second	third
Shape Context [14]	25/25	24/25	22/25
Gdalyahu et. al. (1999)	25/25	21/25	19/25
Sharvit et. al. (1998) [118]	23/25	21/25	20/25
IDSC+DP [121]	25/25	24/25	25/25
Proposed Method	25/25	24/25	24/25

4.2.4 Articulated Tool Data Set

FIGURE 4.5: Some results for the articulated tool database are given. The images on the left are query images, while the four nearest neighbors are shown on their right for the following tools: (top-down) (a) scissor(1), (b) plier(2), (c) knife, (d) plier(1), (e) pincer, (f) scissor(2), plier(3)

In order to evaluate the performance of the proposed algorithm under articulation, we test it using the articulated tool data set [119]. The data set contains 35 shapes: 5×2 scissors, 5×3 pliers, 5 pincers, 5 knives. Each shape differs by an articulation. Some of the results are shown for each tool object in Fig. 4.5. The results for the entire data set are summarized in Table 4.4. The results show that the proposed algorithm is 100 percent accurate for the first nearest neighbor, and its accuracy decreases gradually. Performance for both scissor (1) and scissor

(2), plier (2) and knife retrieval is very good with only one incorrect match in the four nearest matches retrieved. Performance for plier (1) and plier (3) decreases with the fourth nearest match, but most incorrect matches are within the plier sub-group. Pincer is the only tool which is hard to detect mainly because of its similarity with pliers.

TABLE 4.4: Retrieval accuracy for the Articulated Tool data set is given. All the first matches are correct. The accuracy decreases for the fourth nearest neighbor, especially for plier (1), plier (3) and pincer.

tool	Nearest Neighbor			
	first	second	third	fourth
scissor (1)	5/5	5/5	5/5	5/5
scissor (2)	5/5	5/5	5/5	4/5
pincer	5/5	4/5	3/5	1/5
knife	5/5	5/5	5/5	5/5
plier (1)	5/5	5/5	5/5	1/5
plier (2)	5/5	5/5	5/5	5/5
plier (3)	5/5	5/5	3/5	1/5
total	35/35	34/35	31/35	22/35

4.2.5 MNIST Handwritten Digit Database

FIGURE 4.6: Percentage classification error for the MNIST database. Different training set sizes were used to calculate the error. Since we use a posterior probability to classify the object according to the nearest neighbor matches, the error percentages were calculated for different k nearest neighbors. The lowest error was obtained for $k = 15$ and training set size of 10000.

Our shape matching algorithm was tested using the MNIST Digit Database [120], [123]. The database consists of handwritten digits with various degrees of non-linear natural deformations. A training set of 60,000 digits and a test set of 10,000 digits are available where each digit is an 8-bit 28×28 pixel gray image.

The MNIST dataset due to its larger training set makes it easier for machine learning algorithms to find matching data. A larger neighborhood could be searched and used for classification

purposes which was not possible with any of the kimia datasets. For large data sets like MNIST, we incorporate a posterior probability ρ . An image was labeled based on the highest number of class labels present in the k nearest matches. The human error rate calculated for this database was 0.2 percent which could be considered a lower bound for any recognition algorithm being tested on this database. When calculating the mismatch error we have considered digit 6 and 9 as the same digit as they are the rotated versions of each other.

TABLE 4.5: Error performance of different algorithms compared with our proposed method on MNIST digit database.

Algorithm	Error Rate
Lauer et. al. (2007)	0.54
Salakhutdinov and Hinton (2007)	1
Kegl et. al. (2009)	0.87
Ciresan et. al. (2010)	0.35
Jarrett et. al. (2009)	0.53
Ranzato et. al. (2006)	0.39
Proposed	0.39

A number of studies have been done using this database. The error performance for different algorithms has been calculated for the MNIST database and the results are given in [124] and the MNIST database website [123]. The results are reproduced in Table 4.5 for comparison purposes. Deep convolution networks [125] currently gives the best performance on the MNIST database. These neural network algorithms require extensive training and a large feature set to reach an accuracy of 0.35 percent error rate. On the other hand the proposed algorithm requires no training but searches a feature space for nearest matches directly and only uses the training set to calculate a posterior probability based on nearest matches. It is capable of achieving good accuracy even with a relatively small training set. Experimental results for the proposed algorithm using MNIST database are shown in Fig. 4.6. Even with a training set of 100 sample images, the proposed algorithm was able to produce an error rate of 0.63 percent. The best performance, 0.39 percent error rate, was achieved for $k = 15$ and a training set size of 10,000. An error rate of 0.48 percent was achieved with only 1000 training images, but the speedup was significant. This is better than most of the studies done on the database and only slightly worse than the deep convolution neural nets of Ciresan et. al. [125] which provides the best performance.

A lot of the errors stemmed from the outlier removal step where certain shapes would change if only 10 percent of the interest points were removed (e.g. 6 and 8, 1 and 7). There is still room for further improvement in accuracy by changing algorithm parameters such as the number of nearest neighbors in feature extraction and the number of scales used. Increasing the training set size beyond 10,000 did not produce better results. Some specific digits are hard to differentiate and most of the source of mismatches were related to them. Digits like 8 and 3, 9 and 4, or 5 and 6, were harder to differentiate even when a light deformation was present. On the other hand

there were other digits (1, 7, 2 and 0) which were a lot easier to differentiate. Misclassification rates for each digit are shown in Table. 4.6.

TABLE 4.6: Classification performance of the proposed algorithm on MNIST database using a training set size of 10,000 digits

digit	No. of incorrectly classified digits		
	$k = 5$	$k = 10$	$k = 15$
0 zero	2	2	1
1 one	3	3	0
2 two	7	3	3
3 three	6	7	6
4 four	12	9	7
5 five	8	5	5
6 six	8	7	6
7 seven	9	4	2
8 eight	6	4	5
9 nine	7	4	4

4.3 Computation Time

Computational efficiency is dependent on two factors, one is dimensionality and the other is matching procedure [111]. Computational bottlenecks associated with extracting boundaries and aligning objects are described in [110]. Low order moments are computationally less expensive [66]. We use moments below the 5th order which are also more robust to data variation and outliers. Computational complexity for the algorithm is dependent on the complexity of the k-nearest neighbor implementation. Our proposed algorithm is broken down into two different parts, calculating shape descriptor and point matching.

Our algorithm was implemented on an Intel Core i5 CPU, with 2.27 GHz clock speed and 4 GB RAM. It took 150 milliseconds to calculate feature vectors for two images, alignment and then similarity calculation. This also included boundary point extraction (1000 interest points per image). This is better than a previous technique [14] which took 200 ms to match two images. The technique in [14] used 100 sample points where as our sample points generally varied between 1000 and 2000 and our dependent on the image sizes. Higher sampling rate roughly translates into better classification. We reran the experiment using only 500 sample points. This time it took only 50 milliseconds to complete one match.

4.4 Conclusion

Experimental studies were performed for the new point-based shape descriptor described in Chapter 3. It was experimentally shown to have high descriptive power, and computation speed.

It was also found to be robust to complex deformations, articulation, occlusions and missing points.

Chapter 5

Point-based Non-rigid Registration of Medical Images

In this chapter, we present the extensions of point-based descriptors to medical image registration.

5.1 Deformable Rigid Body Hausdorff Registration for Multi-modal Medical Images

Our main contribution in this work is that we implement a multi-scale sub-division registration scheme using the Hausdorff similarity measure combined with differential geometry based feature extraction schemes to obtain meaningful features for registration. These features are based on principal curvatures of the 3D MRI and CT images. Our approach also extends the Hausdorff similarity measure to fit the sub-division volume registration method described in [126]. A similar approach in [127] implements Hausdorff similarity measure combined with the b-splines method for nonrigid registration. B-splines and more general free form deformation methods suffer due to visual and anatomical differences between multi-modal images.

5.1.1 Differential Geometry-based Feature Extraction for Image Registration

Image registration could be based on intensity-based, feature-based or object-based similarity measures. Object based similarity measures require the definition or segmentation of objects in the image before registration which is computationally intensive. Intensity-based similarity measures generally depend on image histograms or entropies which could be quite inconsistent

for multi-modal images. Feature-based similarity measure generally prove to be the most important methods for multi-modal images [128]. Feature extraction is therefore a very important first step, which helps us to locate relevant features which act as landmark or reference points for our image registration step. Different features such as arteries, veins or bony structures are important features that are very useful in the registration step. It must be noted that the features used should be invariant when seen from any imaging technology, viewing angle or deformation. Many ridge, edge and corner detectors exist which could be used for extracting relevant features but these features are generally noisy and may detect irrelevant structures which may not be helpful for registration.

5.1.1.1 Differential Geometry - An Overview

To seek robust and invariant features we use differential geometry based methods [129] [130], which have frequently been used in medical imaging to extract meaningful information from medical images. These methods have widely been used in medical diagnostics, therapy planning and surgery [131], [132]. Differential geometry is helpful in exploiting the curvature and crestlines of the surface of anatomical and physiological structures. In liver surgery, differential geometry based methods could be helpful in identifying blood vessels which could help surgeons plan a surgical intervention [133]. Many fast neural network based implementations are currently available for extracting principal curvatures [134] [135], which makes this feature extraction technique more feasible in a real-time clinical environment. It also reduces the computational load for the Hausdorff similarity measure as we would have a smaller but much more meaningful feature space.

5.1.1.2 Finding Crest Points through Principal Curvatures

A 3D surface could be described by two principal curvatures and their corresponding orthogonal principal directions. These features are also invariant under affine transformations. Since we are treating deformation as local rigid body registration which consists of translations and rotations hence the features chosen are more robust for our sub-division method. The principal curvatures correspond to the eigenvalues of the Weingarten matrix (5.3) and the principal directions are the eigenvectors. The 3D surface is thus transformed into a R^4 hyperspace and the local maxima points of this hyperspace are called the crest points [136]. In essence a principal curvature describes how a surface bends in different directions.

$$\phi_1 = \begin{pmatrix} 1 + I_x^2 & I_x I_y & I_x I_z \\ I_x I_y & 1 + I_y^2 & I_y I_z \\ I_x I_z & I_y I_z & 1 + I_z^2 \end{pmatrix} \quad (5.1)$$

$$\phi_2 = \frac{1}{\sqrt{D}} \begin{pmatrix} I_x I_x & I_x I_y & I_x I_z \\ I_x I_y & I_y I_y & I_y I_z \\ I_x I_z & I_y I_z & I_z I_z \end{pmatrix} \quad (5.2)$$

$$W = \phi_1^{-1} \phi_2 \quad (5.3)$$

ϕ_1 and ϕ_2 are the first and second fundamental forms which are defined in the tangent plane of the surface by the first and second derivatives in x , y and z directions, given by I_x , I_y and I_z . The Weingarten matrix W is obtained in (5.3). Three eigenvalues and eigenvectors are obtained from the Weingarten matrix, corresponding to three principal curvatures and their directions. The principal curvatures are ordered as maximal, medium and minimal curvature. The maximum curvature and its corresponding eigenvector helps us in finding the most meaningful features in the 3D space [136]. A point in the 3D space is selected as a crest point if the maximum curvature c is locally maximum in the curvature eigenvector direction d .

To reduce noise only those crest points are selected that are connected to a certain number of other crest points. Isolated crest points that are not connected to any other crest point or that are connected to only a few crest points are ignored. This way crest points join to form crest lines as shown in Fig. 5.1 and a lot of unwanted noisy features are removed. We use multi-scale techniques to separate the obtained features according to their resolution and importance which will then be used for sub-division registration.

5.1.1.3 Multi-scale Analysis and Feature Extraction

Different features and landmarks in medical images exist at different scales. Medical images when seen from coarse scales tend to show only the larger and more important features. Underlying structures and details in a medical image become more and more visible as you move from coarse to finer scales. Reference. [137] describes the advantages of using the multi-scale hierarchical approach for medical image registration. As we are using a sub-division method, therefore we will be analyzing and registering images at different resolutions. Multi-scale analysis will help us reduce or increase our feature space according to the resolution of our image block. Noise in multi-modal images could reduce the accuracy of our registration scheme. Unwanted noisy features exist mostly at very fine scale which represents the high frequency content. Noise generally gets reduced at coarse scales which represents the low frequency content of the image. We implement multi-scale technique by convolving a 3-dimensional Gaussian kernel G with an input image $L(x, y, z)$ shown in (5.5).

FIGURE 5.1: The MR (top-left) and CT (top-right) image slices are shown. The bottom row shows the crest lines obtained for MR (bottom-left) and CT (bottom-right) by using principal curvatures. The crestlines represent important features in both images which will be used in our registration scheme

$$G(x, y, z; s) = \frac{1}{2\pi s} e^{-(x^2+y^2+z^2)/2s} \quad (5.4)$$

at a certain scale s to provide a scale-space representation.

$$L(x, y, z; s) = G(x, y, z; s) * L(x, y, z) \quad (5.5)$$

The scale s is chosen differently for different multi-modal images depending on the anatomical structures being imaged, as feature sizes vary significantly depending on the distance of the patient to the imaging device. If an image consists of small structures, then multi-scale analysis is done at finer resolutions compared to images which have large structures. Crest lines are then obtained for each of these multi-scale images in a hierarchical manner. Sub-division registration technique first focuses on registering larger and more global features and then focuses on the registration of finer structures. The next section describes the 3-D volume sub-division method employed for registration.

5.1.2 3D Rigid Body Registration-based on the Hausdorff Similarity Measure

FIGURE 5.2: Image sub-division shown, CT image is successively broken down into smaller and smaller blocks and each block is individually registered by translating it over the MR image until a closest match is obtained determined by the Hausdorff similarity measure

Elastic deformation is modeled as a series of local rigid body translations and rotations [138]. The volume sub-division based registration model has been used frequently to model elastic deformation. We use a hierarchical octree based volume sub-division. A multi-modal 3D volume is first broken down into a series of smaller volumes and each smaller volume is then locally registered. Fig. 5.2 shows a 2D description of image sub-division and this method is extended to 3D. These individual volumes are rotated and translated to obtain the closest match. The sub-volume translation should traverse less than half the sub-volume size, otherwise it will then intrude into the neighboring sub-volume. As we are using a feature-based registration scheme hence we only need to divide our feature space which consists of crest line coordinates. After the translation and rotation parameters are obtained, the entire image is then reconstructed using various interpolation scheme. An image sub-division scheme is considered to be the fastest technique for solving image registration problems [139].

5.1.2.1 Multi-scale Feature Sub-division

We have already stressed upon reducing the size of our feature space to reduce computation. Volume sub-division iteratively breaks the volume into quartets. Since each step reduces the size of the image volume to smaller and smaller sub-volumes therefore it inherently makes image registration a multi-scale problem. Reference [126] which follows a sub-division registration method uses a Normalized Mutual Information similarity measure. This measure degrades at smaller volumes and reduces accuracy. Looking at images at the correct resolution would improve accuracy as unwanted noisy features would be eliminated. Accuracy at lower sub-volumes could be increased if a method could be devised which could keep the number of feature points in the

smaller and larger sub-volumes equal. Hausdorff distance becomes computationally expensive due to the large number of features present at larger volumes and degrades for smaller sub-volumes. Multi-scale helps to reduce features for the larger volumes by using larger scales, and increases features by reducing scale as sub-volume size decreases. Multi-scale matching therefore becomes more computationally efficient and accurate as only the relevant features present at a relevant scale are matched.

Since an image volume is only represented by a few pixels which are identified as important features at that particular scale, therefore we only need to know the orientation of these points relative to the center of the block and their location. The problem becomes much simpler as we only have to rotate and translate a vector containing the relevant locations of the pixels and not the entire 3-D volume. The following equations describe the rotation and translation of a 2-D feature vector F to get the new feature vector R . Angle of rotation is given by θ and the translation parameters are T_x and T_y .

$$F = \begin{pmatrix} x_1 + T_x & x_2 + T_x & \cdots & x_n + T_x \\ y_1 + T_y & y_2 + T_y & \cdots & y_n + T_y \\ z_1 + T_z & z_2 + T_z & \cdots & z_n + T_z \end{pmatrix} \quad (5.6)$$

$$RotationMatrix = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (5.7)$$

5.1.2.2 Conventional Hausdorff Distance

The Hausdorff distance has previously been used as a similarity measure in medical image registration and object matching. The Hausdorff distance between point sets $A = \{a_1, a_2, \dots, a_m\}$ and $B = \{b_1, b_2, \dots, b_n\}$ is given by $H(A, B)$:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (5.8)$$

where the directed Hausdorff distance $h(A, B)$ is given by:

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (5.9)$$

The directed Hausdorff distance $h(A, B)$ finds the maximum distance of the set A to the nearest point in set B . Distance between a point in A is first calculated with all the points in B and the smallest distance is selected. This is done for all the points in A and the largest distance value obtained is selected. In a similar manner $h(B, A)$ is calculated.

FIGURE 5.3: Test images shown in the top two rows, the original image is shown in top-left and the deformed image is shown in top-right, middle row shows the translation parameters obtained for the rows (middle-right) and columns (middle-left) for the feature points, and the bottom row shows the the image obtained after image block translations (bottom-left), and the final image after cubic interpolation (bottom-right)

Hausdorff distance in its original form is not robust to noise and image occlusion and needs to be modified. Partial Hausdorff distance used in [12] is a modified version of the conventional Hausdorff distance which provides more robustness to missing features and noise. In partial Hausdorff distance $H_p(A, B)$, k^{th} ranked maximum value is used instead of using the overall maximum distance for calculating $h(A, B)$.

$$h_p(A, B) = \max_{a \in A, b \in B}^{k^{th}} ||a - b|| \quad (5.10)$$

5.1.2.3 Optimization

Optimization is achieved by comparing the Hausdorff distance of the local sub-volume feature registration H_L , to the Hausdorff distance of the global feature registration H_G . Any translation or rotation in the feature space should result in the optimization measure defined by H_C to be lesser than 1. If H_C is greater than 1 then those translation/rotation parameter should be ignored as that will indicate a deterioration in our image registration. The local volume features are rotated and translated and those translation and rotation points are selected which minimize the global Hausdorff distance.

$$H_C = \frac{H_L}{H_G} \quad (5.11)$$

5.1.3 Interpolation and Reconstruction - Deformation Field

We obtain a displacement vector which represents a 3-D displacement for each of our feature points. Since these displacements are only for a few points in the 3-D space hence these values need to be interpolated for the other points. We therefore interpolate these values for each of the 3 dimensions by using tri-cubic interpolation methods [140] to get 3, 3-D matrices X, Y, Z where each point represents a displacement for the point in the original 3-D image. [126] also discusses the possibility of constructing a deformation field. A 2D deformation field could be seen in Fig. 5.3. We use these deformation fields to re-map each pixel in the original 3-D image $L(x, y)$ to its new location $D(x, y)$ by adding the displacements in (X, Y, Z) to its current coordinates (x, y, z) .

$$D(x, y, z) = L(x + X(x, y, z), y + Y(x, y, z), z + Z(x, y, z)) \quad (5.12)$$

The resulting matrix $D(x, y, z)$ could be sparse with many missing data values due to stretching and deformation caused by (5.23). We therefore perform a second interpolation to get the interpolated values for these empty matrix entries. Two or more pixels in $L(x, y, z)$ could also be translated to the same pixel in $D(x, y, z)$ and their values are averaged out. This deformation method is more elastic compared to some of the sub-volume interpolations [126] where they perform translations and rotations on an entire sub-volume. By using just a few feature points in a sub-volume and then creating deformation fields makes sub-volume translations less rigid as could be seen in Fig. 5.3.

At times, image sizes and their resolutions obtained from different imaging technologies could vary significantly. As we have discussed earlier if two pixels are translated to the same location then their values are averaged out. This loses information, and to avoid such cases we should try to register smaller images to the larger images. This would guarantee that the image would have to be stretched for registration and fewer such cases will emerge where we would lose information.

5.1.4 Experiments

We run our algorithm on a publicly available multi-modal database [141]. The database contains 3-D MRI and CT images. The resolution of the MRI images is $100 \times 256 \times 256$ and for CT is $320 \times 256 \times 256$. Multi-scale crest lines are obtained for each of these images using the

FIGURE 5.4: Rigid-body Hausdorff registration system

FIGURE 5.6: The ratio of Hausdorff distances for crestlines is obtained after and before registration for the first 5 images of goldcrest database. The results show that the Hausdorff distance after registration decreases significantly.

algorithm described in Section II. Relevant features are extracted and then sent to sub-division Hausdorff registration system shown in Fig. 5.4. After extracting the displacement vectors for our feature set the final result is interpolated.

To measure the accuracy for registration, [126] uses a simple subtraction scheme which measures the energy content of the difference image obtained after subtracting the original image from the registered image. Since multi-modal images have different structures and intensities present hence a simple difference image may not provide a good measure for the accuracy of registration. We use differential geometry to obtain principal curvatures of the images, which highlight the important landmarks in the original and registered image. Hausdorff distance is then measured between the CT and MRI features before and after registration. A simple ratio is obtained shown in Fig. 5.6 comparing Hausdorff distances before and after registration. Overlapping features (crest lines) will produce a result close to zero for partial Hausdorff distance. The results in Fig. 5.6 show that there is a considerable decrease in Hausdorff distance by almost 5 times.

5.2 Feature Based Nonrigid Registration using Quaternion Sub-division

A new weighted quaternion based nonrigid registration is presented. Strong crest points derived from principal curvatures provide the most robust features for image registration. Crest point strengths are based on their principal curvatures and the number of scales a particular crest point is detected at. Geometric features are extracted which are invariant to rotation, translation and scaling by using neighborhood crest points only as other voxels are susceptible to deformation. The neighborhood size is adjusted according to scale adaptively using a fixed k nearest neighbor

FIGURE 5.7: (a) shows principal curvatures extracted from Brain MRI with $\sigma = 2$, (b) shows Crestlines extracted from Brain MRI with $\sigma = 2$.

to make the extracted feature scale invariant. Statistical properties are used to measure the distribution of these geometric invariant features. The scale and rotation invariant feature points are then used to establish a point to point correspondence between the template crest points and the subject image crest points. A multi-scale feature-based sub-division scheme is employed for registration where a weighted quaternion matrix provides a quaternion transformation based on the corresponding points to obtain the best rotation for global as well as local sub-blocks.

5.2.1 Feature Point Extraction

Feature extraction is a very important first step, which helps us to locate relevant features which are treated as landmark or reference points for our image registration step. Different features such as arteries, veins, bony structures or organ surface boundaries are important reference points that are very useful in the registration step. It must be noted that the features used should be invariant when seen from any imaging technology, viewing angle or deformation. Sub-division registration schemes generally use windowing methods to extract invariant features [126]. Different types of windows and different statistical invariants are used. Work has already been done on many geometric and algebraic invariants [115], [116]. The window used is again a limiting factor since it is a rigid window which cannot adaptively change according to the complex deformation of the underlying image. This introduces an error as statistical information would be affected if the sub-block is deformed. Also, if statistical information is not limited to important and more registration relevant features then irrelevant and redundant features might affect the accuracy of the description. If important features move out of a window then accuracy of the invariant feature descriptors for a particular region would change drastically.

Based on the above limitation we deviate from the common window based similarity techniques and propose a two step feature extraction step. In the first step we locate interest points which

would be robust to nonrigid deformation and in the second step we extract relevant rotation and scale invariant geometric features from those interest points only. We use differential geometry [129] to extract surface features of anatomical and physiological structures. Similarly crest points could be used to identify ridges and valleys on iso-surfaces in brain MRI scans. Many fast implementations are currently available for extracting principal curvatures which also incorporate multi-scale [134].

A 3D surface is described by two principal curvatures and their corresponding orthogonal principal directions. A 3D surface or boundary region separating different regions in 3D based on their intensity form iso-surfaces and the zero-crossing point of these iso-surfaces are described by crest points. These features are not invariant under affine transformations so to improve robustness we reject crest points which are below a certain threshold under the assumption that high curvature crest points would survive deformations better than low curvature crest points. Crest point also allows us to reduce the huge amount of 3D data and make registration more efficient by focusing on a limited number of interest points. Three eigenvalues and eigenvectors are obtained from the Weingarten matrix [129], corresponding to three principal curvatures and their directions. The principal curvatures are ordered as maximal, medium and minimal curvature. A point in the 3D space is selected as a crest point if the maximum curvature c is locally maximum in the curvature eigenvector direction d . To reduce noise only those crest points are selected that are connected to a certain number of other crest points. Isolated crest points that are not connected to any other crest point or that are connected to only a few crest points are ignored. The crest points join to form crest lines and unwanted noisy features are hence removed. We use multi-scale techniques to separate the features obtained according to their resolution and importance which will then be used for sub-division registration. The multi-scale principal curvature values for each voxel and the crest points obtained after thresholding are shown in Fig. 5.7.

5.2.1.1 Geometric Invariant Feature Extraction

Crest points need to be uniquely identified hence further rotation and scale invariant features should be extracted to differentiate between crest points. Since crest points are the most robustly identifiable feature in a deformed image hence any invariant feature must be extracted using crest points only. We consider the remaining voxels susceptible to deformation variation. Geometric invariant primitives [116] which consist of distance D between two arbitrary points, the area A between three arbitrary points, the dot product d_p between two vectors $\vec{p}_i$ and $\vec{p}_j$ are invariant to rotation and translation. The dot product is the only property which is invariant to scale as well. Area A and distance D can be made invariant to scale by normalizing them to 1.

FIGURE 5.8: Crest point surrounded by five nearest neighbor crest points is shown, where $\{d1, d2, d3, d4, d5\}$ are the distances between five nearest neighbor points with the crest point. Arrows represent the direction of principal curvature for the five nearest neighbor crest points.

$$D(x_i, x_j) = \sqrt{(x_i^1 - x_j^1)^2 + (x_i^2 - x_j^2)^2 + (x_i^3 - x_j^3)^2} \quad (5.13)$$

$$A(\vec{p}_i, \vec{p}_j) = \frac{1}{2} \|\vec{p}_i \times \vec{p}_j\| \quad (5.14)$$

$$d_p = \vec{p}_i \cdot \vec{p}_j \quad (5.15)$$

$$\theta_i = \cos^{-1} \left(\frac{\vec{p}_i \cdot \vec{p}_j}{\|\vec{p}_i\| \|\vec{p}_j\|} \right) \quad \text{where } j = 1, 2, \dots, k \quad (5.16)$$

$$\theta_i^k = [\theta_{i1}, \theta_{i2}, \dots, \theta_{ik}] \quad (5.17)$$

$$D_i^k = [D_{i1}, D_{i2}, \dots, D_{ik}] \quad (5.18)$$

Since we limit our invariant features to crest points hence a k nearest neighbor search provides us with k nearest crest points for each interest point in the image. Principal curvature eigenvectors $\vec{p}$ are taken as feature points and dot products are calculated between each k nearest neighbor $\vec{p}_j$ with the crest point principal curvature eigenvector $\vec{p}_i$. This gives us a set θ_k consisting of angles which the principal curvature eigen-vector makes with the surrounding crest point neighborhood as shown in Fig. 5.8. Moment invariant features are then extracted from this set θ_k consisting of mean μ , variance σ^2 and kurtosis values μ^4/σ^4 (or fourth standardized moment) which represent low order moment invariants which are invariant to rotation and translation. Similarly moments are calculated for principal curvature eigenvalues e_1 , e_2 and e_3 (maximal, medium and minimal direction) for k nearest neighbor crest points, which are also robust to rotation, translation and scaling. A normalized distance distribution D_i^k is also generated between k nearest neighbor crest points and low order moments are extracted from the distribution. These features would help us determine the similarity for point to point correspondence between the subject and template image.

The k nearest neighbors are fixed quantities and they should be sufficiently large to meet statistical requirements. Since multi-scale features would be more spread out at larger scales and more compactly located at finer scales hence a fixed value of k would automatically increase or decrease the size of the region from which neighboring crest points are extracted. The advantage of using such a scheme involving k nearest neighbors is that there is no fixed window being used. Some areas would have more features compared to other areas where features would be sparse, hence k nearest neighbor would automatically adapt the area being covered. The limitation associated with area based methods especially for regions where features are less is hence removed. A second advantage would be that feature matching is not limited by rigid sub-blocks. If a feature point in one sub-block of the template image matches with a feature from another sub-block of the subject image then these points are paired. Such a situation might occur due to deformation.

5.2.2 Registration Using Quaternions

FIGURE 5.9: (a) This graph represents the effect of adding noise to the performance of quaternion registration, (b) This graph represents the effect of increasing the number of non corresponding features on the performance of quaternion registration. Less similar or non-corresponding features are assigned weights $w(x, y, z)$ to increase the accuracy of the registration process.

Quaternion fitting techniques based on the Iterative Closest Point (ICP) algorithm provides a closed form solution for a feature-based registration problem [10]. Quaternion fitting techniques provide the closest match between corresponding registration points in R^3 in a least squares sense. Successive shifting and rotation is eliminated during sub-division registration which saves computation cost. We choose quaternion based ICP algorithm as quaternion multiplication takes sixteen multiplications and twelve additions, compared with orthogonal rotation matrices [142] which require twenty-seven multiplications and eighteen additions.

A coordinate (x_n, y_n, z_n) for a feature point p_n in R^3 is represented as the imaginary part $\vec{v}_n$ in quaternion space with real part equal to zero. A quaternion with real part equal to zero is called a pure quaternion. $L_q(v)$ rotates the vector $\vec{v}_n$ through an angle θ about $\vec{q}$ as the axis of rotation

where $\vec{q}$ is the imaginary part of q . The imaginary part of pure quaternion $\vec{v}_n$ is represented by vector $\vec{v}_n$. Unit quaternion operator q does not change the length of the vector $\vec{v}_n$ and if the direction of q is along $\vec{v}_n$ then it is left unchanged. Quaternion rotation implementation has been done using matrices as shown in Eq. (5.20).

$$L_q(v) = q\vec{v}_nq^* \quad (5.19)$$

$$q\vec{v}_nq^* = \begin{pmatrix} q_1 & -q_2 & -q_3 & -q_4 \\ q_2 & q_1 & -q_4 & q_3 \\ q_3 & q_4 & q_1 & -q_2 \\ q_4 & -q_3 & q_2 & q_1 \end{pmatrix} \begin{pmatrix} q_1 & -q_2 & -q_3 & -q_4 \\ q_2 & q_1 & q_4 & -q_3 \\ q_3 & -q_4 & q_1 & q_2 \\ q_4 & q_3 & -q_2 & q_1 \end{pmatrix} \begin{pmatrix} 0 \\ x_n \\ y_n \\ z_n \end{pmatrix} \quad (5.20)$$

Faugeras and Hebert [143] and Horn [10] have proposed a quaternion registration technique to register a set of points $S^1 = \{P_i^2\}$ with a another corresponding set of points $S^2 = \{P_i^1\}$. The condition for registration is that each point must correspond precisely with the same point in the rotated and noisy version. This would present a problem during registration of normal images as feature points extracted are not indexed and a one to one correspondence needs to be extracted. This sort of correspondence could be extracted by finding similarity in a rotation, translation and scale invariant feature space.

The unit quaternion representing the best rotation is obtained by extracting the eigenvector of the 4×4 matrix given in Eq. (5.21).

$$\begin{pmatrix} S_{xx} + S_{yy} + S_{zz} & S_{yz} - S_{zy} & S_{zx} - S_{xz} & S_{xy} - S_{yx} \\ S_{yz} - S_{zy} & S_{xx} - S_{yy} - S_{zz} & S_{xx} + S_{xx} & S_{xx} + S_{xx} \\ S_{xx} - S_{xx} & S_{xx} + S_{xx} & -S_{xx} + S_{yy} - S_{zz} & S_{xx} + S_{xx} \\ S_{xx} - S_{xx} & S_{xx} + S_{xx} & S_{xx} + S_{xx} & -S_{xx} - S_{yy} + S_{zz} \end{pmatrix} \quad (5.21)$$

where $S_{aa} = \sum_{i=1}^n \hat{a}_i^2 \hat{a}_i^1$ where $\hat{a}_i^j = \hat{x}_i^j, \hat{y}_i^j$, or $\hat{z}_i^j$ and $\hat{x}_i^j, \hat{y}_i^j, \hat{z}_i^j$ being the centered or mean subtracted coordinates of the points P_i^j . We have simplified translation by simply subtracting the mean from the subject image feature points and centered it on the template image point. A weight $w(x, y, z)$ is introduce for each feature point. This weight corresponds to the degree of mismatch between feature points and the number of scales that a feature point is present at. Experiments were conducted to analyze the performance of a weighted quaternion fitting technique [74]. A rigid registration was carried out in an artificial set of data where the number of non-corresponding points were increased. Since both template image and subject image undergo different types of deformation hence a number of feature points with no correspondence exist. If

template feature points are searched in the subject feature space then a number of close matches would be found which may not be a perfect match. A degree of uncertainty would exist between the two matched points based on the euclidean distance similarity. Threshold would then be applied to each point based on the importance of the feature points (detection scale) and its similarity with its corresponding point. Fig. 5.9b shows a comparison of different weights assigned to different number of non-corresponding points. Hausdorff similarity measure is then used to assess the registration result. It is clear from Fig. 5.9b that as the weight for non-corresponding points decreases, then registration results improve. We would adapt this technique for our case as well and apply a threshold which will determine the degree of mismatch between corresponding points. The weights used would be adaptable. The fitting quaternion obtained from Eq. (5.21) would have to obey a global limitation for the angle by which it rotates a particular feature sub-block. This feature sub-block must rotate within a pre-described limits of $\pi/4$ radians. Any rotation beyond that would lead to a deformation which would be unrealistic for a real case. For weighted quaternion fitting S_{aa} could be rewritten as shown in Eq. (5.22).

$$S_{aa} = \sum_{i=1}^n w(i) \hat{a}_i^2 \hat{a}_i^1 \quad \text{where} \quad \hat{a}_i^j = x_i^j, y_i^j \text{ or } z_i^j \quad (5.22)$$

5.2.2.1 Sub-division Registration Using Unit Quaternion

The sub-division registration scheme differs from previous sub-division schemes as it only partitions feature points according to the 3D space but does not create rigid sub-volumes. The shifting and rotation performed would be limited to the feature points. Since feature points are few in number hence rotation and shifting would be much faster. We also use a multi-scale approach which would limit the registration step to important features at higher scales and reduce the number of feature points by removing finer scale descriptors. As sub-division proceeds to smaller sub-blocks, more and more fine scale feature points are introduced to increase the accuracy of registration of smaller blocks. A 3D Gaussian kernel is used with standard deviation $\sigma = 1, 2$ and 4 .

5.2.3 Deformation Field and Interpolation

We obtain a displacement vector after registration which represents a 3-D displacement for each of our feature points. Since these displacements are only for a few selected feature points in the 3-D space hence these values need to be interpolated for the other points. We therefore interpolate these values for each of the 3 dimensions by using tri-cubic interpolation methods [140] to get 3, 3-D matrices X_d, Y_d, Z_d where each point represents a displacement for the point in the original 3-D image. A similar deformation field is constructed in [126]. We use

these deformation fields to re-map each pixel in the original 3-D image $L_{deformed}(x, y, z)$ to its new location $L_{registered}(\bar{x}, \bar{y}, \bar{z})$ by adding the displacements in (X_d, Y_d, Z_d) to its current coordinates (x, y, z) .

$$L_{registered}(\bar{x}, \bar{y}, \bar{z}) = L_{deformed}(x + X_d(x, y, z), y + Y_d(x, y, z), z + Z_d(x, y, z)) \quad (5.23)$$

The resulting matrix $L_{registered}(x, y, z)$ could be sparse with many missing data values due to stretching and deformation caused by (5.23), hence a second interpolation is performed after relocating data points to get these values. Gaussian smoothing windows are not necessary for smoothing out boundary shifts as the first interpolation provides a smooth boundary between sub-blocks. By using just a few feature points in a sub-volume and then creating deformation fields makes sub-volume translations less rigid with lesser discontinuities as could be seen in Fig. 5.10c.

5.2.4 Experiments

We use brain scans from the Internet Brain Segmentation Repository (IBSR). The data is provided by the Center for Morphometric Analysis at Massachusetts General Hospital and is available at “<http://www.cma.mgh.harvard.edu/ibsr/>”. The experiment is divided into two steps. In the first step we deform MR images by a random deformation field. A deformed image is shown in Fig. 5.10b. In the second step we register and quantitatively measure the registration of this deformed Brain MR to the original MR under varying deformations.

A uniform random deformation field is created with random shifts for each feature vector. The deformation is controlled by changing the maximum number of shifts possible for each pixel. By increasing shifts we increase the random deformation which takes place and then try to register image. The random shifts are stored in X_d, Y_d and Z_d . Similarity is measured using Hausdorff distance to quantitatively measure the performance of our registration scheme to different strengths of deformation. Fig. 5.10d shows the performance and robustness of our algorithm under different deformation strengths compared with a popular registration software ELASTIX (<http://elastix.isi.uu.nl/>).

5.3 Conclusion

We presented a rigid body deformable registration scheme which is robust to noise, intensity variation and missing features. The registration technique focuses on multi-modal images which

FIGURE 5.10: (a) Original Image, (b) Image deformed using a uniformly distributed deformation vector, (c) Results obtained after registering deformed image (b) with (a), (d) measures the registration accuracy using Hausdorff similarity measure against different deformation strengths where blue line represents the results using our technique and red line represents the results obtained using ELASTIX registration software.

have significant intensity variations and different levels of SNR. A huge amount of feature mismatch is also present between modalities as different modalities visualize different tissue and organs differently. We have analyzed the robustness and sensitivity of our algorithm on CT and MRI images of different quality. Due to the real-time nature of Image Guided Interventions we have also made sure that our technique could be speeded up by dedicated hardware architectures. Different Neural network techniques are available for speeding up the crest line algorithm. Our multi-scale approach on the other hand, reduces the number of features fed into the Hausdorff similarity measure based sub-division registration scheme. The multi-scale nature of the crestlines also reduces the computation for our sub-division algorithm.

In the second section, we presented a robust registration method based on geometric invariant features using multi-scale weighted quaternion sub-division. The geometric invariant features are based on selective crest points which are robust to deformation. The k-nearest neighbor approach combined with multi-scale provides a new adaptive method for collecting features. The size of neighborhood increases if fewer features are present and decreases when features are abundant. This is an improvement over previous techniques where lesser features in some parts of the MR image would result in less robust moment invariants. The quaternion fitting

approach also provides a much quicker way for registering during sub-division registration, and the associated feature weights help remove any outlier effect on the registration result.

Chapter 6

Conclusion and Future Work

6.1 Summary

We have developed a new technique for shape classification and retrieval based on a novel point-based shape descriptor. The proposed shape descriptor is based on low order moments extracted from features related to principal curvatures of neighborhood interest points. The new shape descriptor was used for finding point correspondence between two sets of interest points and shape similarity was calculated using alignment and distance measures. The proposed shape descriptor was able to differentiate between reflection symmetries without a global model incorporated. This shape descriptor was tested on several databases and found to be robust to occlusions, noise, deformations and articulation. When compared to existing ones, the proposed shape descriptor has high descriptive power and computation speed.

Our point-based nonrigid shape matching has been extended to medical image registration. One nonrigid registration technique has been based on the sub-division scheme. Hausdorff distance-based sub-division scheme was developed for nonrigid registration. The Hausdorff distance provides a good similarity metric for registration and the partial Hausdorff distance provides additional robustness against outliers. Another sub-division registration technique was developed using weighted quaternion rotation. This technique was based on finding point correspondences. The weighting scheme was based on a Gaussian smoothing function. The weights were adjusted iteratively based on the displacements of points to minimize outliers. We have analyzed the robustness and sensitivity of our algorithms on CT and MRI images of different quality. When compared to popular ELASTIX registration method, ours give better registration accuracy.

FIGURE 6.1: Different modalities for imaging the human brain are given.

6.2 Possible Extensions and Future Work

Although feature-based registration may solve some of the drawbacks associated with intensity-based registration algorithms but the problem remains far too complicated and challenging as can be seen in Fig. 6.1 where different modalities of the human brain are present. Although there are some functional and anatomical correspondences between some images, but in other images these correspondences are simply non-existent as seen in the figure. Functional consistencies at times cannot be extracted even from point features. A lot depends on how a similarity criterion is defined for a registration problem. Some research has gone into removing unnecessary information from images and to drive the registration process using only similar structures. This too has its limitations as some images would simply be too different and would present very few corresponding structures which will not be enough to drive the registration step.

Research in human vision has shown that the human brain uses high curvature regions to detect and identify objects. An edge image of a human face could be sufficient for identification purposes for humans but what is forgotten is that the human brain is capable of storing past images

as references which is why other image features might be considered redundant whenever a human brain sees a new image. The feature extraction step is computationally expensive and one way to speed up feature-based registration would be to have features extracted or known a priori. Such a scheme would reduce the processing required for feature extraction. We will extend our previous work, especially related to weighted quaternion sub-division scheme for achieving nonrigid registration. We will explore registration further focusing on the following problems:

- A lot of research has gone into defining similarity measures. Similarity could be an ambiguous term as can be seen in Fig. 6.1. The entire framework for registration needs to be rethought as similarity is an adaptable measure which changes spatially. Global invariant features are also frequently used for registration, but images cannot possibly have consistently varying invariant features. Features must therefore vary spatially so that similarity could be better determined. Applying the procrustean logic of limiting the problem to “feature-based” or “intensity-based” methods would not solve the problem. A hybrid method should be investigated.
- Multi-scale optimization is now considered a standard practice for solving registration problems. Multi-scale methods have problems when dealing with local maximas. Graph cuts is one interesting approach which could possibly solve registration while dealing with spatial coherence.
- The introduction of better and more efficient algorithms have contributed far more to speeding up computations. Researchers have started to focus more on hardware acceleration which has historically contributed little in achieving desirable speed ups. We will explore ways to improve speed by reformulating the registration problem.
- Custom hardware for computationally intensive applications is an emerging trend in image processing, especially for medical images where one often has to deal with large three dimensional MRI and CT data sets. The aim is to improve computational time while reducing the cost. General Purpose Computing on Graphics Processing Units (GPGPU) have recently become attractive due to the increasing capability of graphics hardware coupled with the increased programmability and functionality of such GPUs. Accelerating our registration method on GPGPU would be an important future contribution.

List of Publications

1. Fahad Hameed Ahmad, N. Sudha, “Deformable rigid body Hausdorff registration for multi-modal medical images”, *Proceedings of IEEE TENCON*, Singapore, Nov. 2009.
2. Fahad Hameed Ahmad, N. Sudha, Jimmy Liu Jiang, “Feature based nonrigid registration using quaternion subdivision”, *Proceedings of International Conference on Medical Biometrics* (LNCS 6165), Hong Kong, June 2010.
3. F. H. Ahmad and S. Natarajan, “Multi-modal Medical Images Registration Using Differential Geometry and Hausdorff Distance”, *Journal of Intelligent Systems*, Vol. 19, No.4, 2010.
4. Fahad Hameed Ahmad, N. Sudha, “Feature-based Directional Shape Descriptor for Shape Retrieval and Classification”, *Machine Vision and Applications* (under review)

Appendix A

Geometric Primitives

The 2D geometric primitives include distance D , area A and dot product d :

$$D(a, b) = [(x_a - x_b)^2 + (y_a - y_b)^2]^{1/2} \quad (\text{A.1})$$

$$A(a, b, c) = \frac{1}{2} \|(x_b - x_a, y_b - y_a) \times (x_c - x_a, y_c - y_a)\|_2 \quad (\text{A.2})$$

$$d(\vec{ab}, \vec{ac}) = (x_b - x_a, y_b - y_a) \bullet (x_c - x_a, y_c - y_a) \quad (\text{A.3})$$

where $\{x_a, y_a\}$, $\{x_b, y_b\}$ and $\{x_c, y_c\}$ represent three different points in a 2D coordinate space.

Appendix B

Low Order Moments

$$\mu(X) = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{B.1})$$

$$\text{var}(X) = \mathbf{E}[(X - \mu)^2] \quad (\text{B.2})$$

$$\delta(X) = \frac{\mu_3}{\sigma^3} \quad (\text{B.3})$$

$$\gamma(X) = \frac{\mu_4}{\sigma^4} \quad (\text{B.4})$$

where $X = \{x_1, \dots, x_n\}$ is the sample space, $\mathbf{E}$ is the expectation operator, σ is the standard deviation, and μ_3 and μ_4 are the third and fourth moments about the mean. The third standardized moment $\delta(X)$ represents skewness of the distribution where as the fourth standardized moment is kurtosis $\gamma(X)$ and describes the peakedness of the distribution.

Appendix C

Quaternion Fitting

A quaternion with real part equal to zero is called a pure quaternion and is given by v_I^r . A coordinate (x_I^r, y_I^r, z_I^r) for a point $[x_I^r, y_I^r, z_I^r]$ in R^3 is represented as the imaginary part of v_I^r in quaternion space R^4 , with real part equal to zero as shown in (C.1) where $\mathbf{v}_I^r$ represents the imaginary part.

$$v_I^r = 0 + x_I^r \mathbf{i} + y_I^r \mathbf{j} + z_I^r \mathbf{k} \quad (\text{C.1})$$

$$\mathbf{v}_I^r = x_I^r \mathbf{i} + y_I^r \mathbf{j} + z_I^r \mathbf{k} \quad (\text{C.2})$$

TABLE C.1: Quaternion Basis Multiplication

$\times$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

The quaternion rotation operator $L_q(v_I^r)$, given by (C.5), rotates the vector $\mathbf{v}_I^r$ through an angle θ about $\mathbf{q}$ as the axis of rotation, where $\mathbf{v}_I^r$ and $\mathbf{q}$ are the imaginary parts of v_I^r and q . Multiplication of basis elements of quaternion are described in Table C. The unit quaternion operator q does not change the length of the vector $\mathbf{v}_I^r$ and if the direction of $\mathbf{q}$ is along $\mathbf{v}_I^r$ then it is left unchanged.

$$q = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k} \quad (\text{C.3})$$

$$\mathbf{q} = q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k} \quad (\text{C.4})$$

$$L_q(v_I^r) = qv_I^r q^* \quad (\text{C.5})$$

The quaternion q that best represents the rotation to fit two sets of matched data points $M(r)$ is obtained using the symmetric 4×4 matrix $\mathbf{Q}_{MAT}$ (C.6), where $A_{aa} = \sum_{r=1}^n \hat{a}_{I_1}^r \hat{a}_{I_2}^r$ and $\hat{a}_I^r = \hat{x}_I^r, \hat{y}_I^r$, or $\hat{z}_I^r$. Both sets of points, $[x_{I_2}^r, y_{I_2}^r]$ and $[x_{I_2}^r, y_{I_2}^r]$, are centered on their centroid to give $\hat{x}_I^r, \hat{y}_I^r, \hat{z}_I^r$. The rotation quaternion q is the eigenvector of the quaternion matrix $\mathbf{Q}_{MAT}$ (C.6). We implement a 2D version of quaternion fitting by replacing $\hat{z}_I^r = 0$ for all matched points. We have simplified translation by subtracting the mean from the subject image feature points and centered it on the template image point. Quaternion rotation $L_q(v_I^r)$ implementation has been done using matrices and is shown in (C.7). The imaginary part of $L_q(v_I^r)$ gives us the new transformed coordinates $[\bar{x}_{I_2}^r, \bar{y}_{I_2}^r, \bar{z}_{I_2}^r]$. We ignore $\bar{z}_{I_2}^r$ as it is zero for the 2D case.

$$\mathbf{Q}_{MAT} = \begin{pmatrix} A_{xx} + A_{yy} + A_{zz} & A_{yz} - A_{zy} & A_{zx} - A_{xz} & A_{xy} - S_{yx} \\ A_{yz} - A_{zy} & A_{xx} - A_{yy} - A_{zz} & A_{xx} + A_{xx} & A_{xx} + S_{xx} \\ A_{xx} - A_{xx} & A_{xx} + A_{xx} & -A_{xx} + A_{yy} - A_{zz} & A_{xx} + S_{xx} \\ A_{xx} - A_{xx} & A_{xx} + A_{xx} & A_{xx} + A_{xx} & -A_{xx} - A_{yy} + S_{zz} \end{pmatrix} \quad (\text{C.6})$$

$$qv_{I_1}^r q^* = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{pmatrix} \begin{pmatrix} 0 \\ x_I^r \\ y_I^r \\ 0 \end{pmatrix} \quad (\text{C.7})$$

Bibliography

- [1] D. Rueckert. Non-rigid registration: Techniques and applications. *Medical Image Registration*, 13:281–302, 2001.
- [2] N.A. Parra. Rigid and non-rigid point-based medical image registration. *FIU Electronic Theses and Dissertations*, page 127, 2009.
- [3] JB Maintz and M.A. Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998. ISSN 1361-8415.
- [4] B. Heisele, U. Kressel, and W. Ritter. Tracking non-rigid, moving objects based on color cluster flow. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, page 257. Published by the IEEE Computer Society, 1997.
- [5] I. Biederman and G. Ju. Surface versus edge-based determinants of visual recognition* 1. *Cognitive Psychology*, 20(1):38–64, 1988.
- [6] WS Geisler, JS Perry, BJ Super, and DP Gallogly. Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41(6):711–724, 2001. ISSN 0042-6989.
- [7] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. In *Eighth IEEE International Conference on Computer Vision, 2001. ICCV. Proceedings*, volume 1, 2001.
- [8] C. Martinez-Ortiz and J. Zunic. Curvature weighted gradient based shape orientation. *Pattern Recognition*, 43(9):3035–3041, 2010. ISSN 0031-3203.
- [9] WR Crum, T. Hartkens, and DLG Hill. Non-rigid image registration: theory and practice. *British journal of radiology*, 77(Special Issue 2):S140, 2004.
- [10] B.K.P. Horn et al. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.
- [11] L.S. Hibbard and R.A. Hawkins. Objective image alignment for three-dimensional reconstruction of digital autoradiograms. *Journal of neuroscience methods*, 26(1):55–74, 1988. ISSN 0165-0270.

- [12] D.P. Huttenlocher, G.A. Klanderman, and WA Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions on pattern analysis and machine intelligence*, pages 850–863, 1993. ISSN 0162-8828.
- [13] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on pattern analysis and machine intelligence*, pages 239–256, 1992.
- [14] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 509–522, 2002.
- [15] M. Holden. A review of geometric transformations for nonrigid body registration. *Medical Imaging, IEEE Transactions on*, 27(1):111–128, 2007. ISSN 0278-0062.
- [16] S.Y. Park, S.I. Choi, J. Moon, J. Kim, and Y.W. Park. Real-time 3D registration of stereo-vision based range images using GPU. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–6. IEEE, 2010.
- [17] A. Quddus and O. Basir. Wavelet-Based Medical Image Registration for Retrieval Applications. In *2008 International Conference on BioMedical Engineering and Informatics*, pages 301–305. IEEE, 2008.
- [18] D. Rueckert and P. Aljabar. Nonrigid Registration of Medical Images: Theory, Methods, and Applications [Applications Corner]. *Signal Processing Magazine, IEEE*, 27(4):113–119, 2010. ISSN 1053-5888.
- [19] D. Mattes, D.R. Haynor, H. Vesselle, T.K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *Medical Imaging, IEEE Transactions on*, 22(1):120–128, 2003. ISSN 0278-0062.
- [20] M. Staring, S. Klein, and J.P.W. Pluim. A rigidity penalty term for nonrigid registration. *Medical physics*, 34:4098, 2007.
- [21] D. Rueckert, LI Sonoda, C. Hayes, DLG Hill, MO Leach, and DJ Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *Medical Imaging, IEEE Transactions on*, 18(8):712–721, 2002. ISSN 0278-0062.
- [22] T. Rohlfing, C.R. Maurer Jr, D.A. Bluemke, and M.A. Jacobs. Volume-preserving non-rigid registration of MR breast images using free-form deformation with an incompressibility constraint. *Medical Imaging, IEEE Transactions on*, 22(6):730–741, 2003. ISSN 0278-0062.
- [23] M. Foskey, B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman, and S. Joshi. Large deformation three-dimensional image registration in image-guided radiation therapy. *Physics in Medicine and Biology*, 50:5869, 2005.

- [24] X. Pennec, P. Cachier, and N. Ayache. Tracking brain deformations in time sequences of 3D US images. *Pattern Recognition Letters*, 24(4-5):801–813, 2003. ISSN 0167-8655.
- [25] T. Rohlfing, R. Brandt, R. Menzel, C.R. Maurer, et al. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage*, 21(4):1428–1442, 2004. ISSN 1053-8119.
- [26] S. Klein, U.A. van der Heide, I.M. Lips, M. van Vulpen, M. Staring, and J.P.W. Pluim. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Medical physics*, 35:1407, 2008.
- [27] R.A. Heckemann, J.V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1):115–126, 2006. ISSN 1053-8119.
- [28] S. Maheswaran, H. Barjat, S.T. Bate, P. Aljabar, D.L.G. Hill, L. Tilling, N. Upton, M.F. James, J.V. Hajnal, and D. Rueckert. Analysis of serial magnetic resonance images of mouse brains using image registration. *NeuroImage*, 44(3):692–700, 2009. ISSN 1053-8119.
- [29] C. Baillard, P. Hellier, and C. Barillot. Segmentation of brain 3D MR images using level sets and dense registration. *Medical Image Analysis*, 5(3):185–194, 2001. ISSN 1361-8415.
- [30] W.R. Crum, R.I. Scahill, and N.C. Fox. Automated hippocampal segmentation by regional fluid registration of serial MRI: validation and application in Alzheimer’s disease. *Neuroimage*, 13(5):847–855, 2001. ISSN 1053-8119.
- [31] L. Zhang, E.A. Hoffman, and J.M. Reinhardt. Atlas-driven lung lobe segmentation in volumetric X-ray CT images. *Medical Imaging, IEEE Transactions on*, 25(1):1–16, 2005. ISSN 0278-0062.
- [32] T.A. Sundaram and J.C. Gee. Towards a model of lung biomechanics: pulmonary kinematics via registration of serial lung images. *Medical image analysis*, 9(6):524–537, 2005. ISSN 1361-8415.
- [33] H.A. Vrooman, C.A. Cocosco, F. van der Lijn, R. Stokking, M.A. Ikram, M.W. Vernooij, M. Breteler, and W.J. Niessen. Multi-spectral brain tissue segmentation using automatically trained k-Nearest-Neighbor classification. *NeuroImage*, 37(1):71–81, 2007. ISSN 1053-8119.
- [34] P. Anbeek, K.L. Vincken, M.J.P. van Osch, R.H.C. Bisschops, and J. van der Grond. Automatic segmentation of different-sized white matter lesions by voxel probability estimation. *Medical image analysis*, 8(3):205–215, 2004. ISSN 1361-8415.

- [35] T. Netsch, P. Roesch, J. Weese, A. van Muiswinkel, and P. Desmedt. Grey-value-based 3D registration of functional MRI time-series: comparison of interpolation order and similarity measure. In *Proceedings of SPIE*, volume 3979, page 1148, 2000.
- [36] L. Freire and J.F. Mangin. Motion correction algorithms may create spurious brain activations in the absence of subject motion. *NeuroImage*, 14(3):709–722, 2001. ISSN 1053-8119.
- [37] L. Bergmans, J. Van Cleynenbreugel, M. Wevers, and P. Lambrechts. A methodology for quantitative evaluation of root canal instrumentation using microcomputed tomography. *International Endodontic Journal*, 34(5):390–398, 2001. ISSN 1365-2591.
- [38] J. Van Cleynenbreugel, F. Schutyser, J. Goffin, K. Van Brussel, and P. Suetens. Image-based planning and validation of C1–C2 transarticular screw fixation using personalized drill guides. *Computer Aided Surgery*, 7(1):41–48, 2002. ISSN 1097-0150.
- [39] C.R. Meyer, J.L. Boes, B. Kim, P.H. Bland, G.L. Lecarpentier, J.B. Fowlkes, M.A. Roubidoux, and P.L. Carson. Semiautomatic registration of volumetric ultrasound scans. *Ultrasound in medicine & biology*, 25(3):339–347, 1999. ISSN 0301-5629.
- [40] V. Zagrodsky, R. Shekhar, and J.F. Cornhill. Multifunction extension of simplex optimization method for mutual information-based registration of ultrasound volumes. In *Proceedings of SPIE*, volume 4322, page 508, 2001.
- [41] N. Ritter, R. Owens, J. Cooper, R.H. Eikelboom, and P.P. Van Saarloos. Registration of stereo and temporal images of the retina. *Medical Imaging, IEEE Transactions on*, 18(5):404–418, 2002. ISSN 0278-0062.
- [42] X.S. Wang, L.G. Li, C.S. Hu, J.J. Qiu, Z.Y. Xu, and Y. Feng. A comparative study of three CT and MRI registration algorithms in nasopharyngeal carcinoma. *Journal of Applied Clinical Medical Physics*, 10(2), 2009. ISSN 1526-9914.
- [43] B. Fischer and J. Modersitzki. Ill-posed medicinean introduction to image registration. *Inverse Problems*, 24:034008, 2008.
- [44] D.L.G. Hill, P.G. Batchelor, M. Holden, and D.J. Hawkes. Medical image registration. *Physics in medicine and biology*, 46:R1, 2001.
- [45] D.L. Collins, P. Neelin, T.M. Peters, and A.C. Evans. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of computer assisted tomography*, 18(2):192, 1994. ISSN 0363-8715.
- [46] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.

- [47] G.K. Rohde, A. Aldroubi, and B.M. Dawant. The adaptive bases algorithm for intensity-based nonrigid image registration. *Medical Imaging, IEEE Transactions on*, 22(11): 1470–1479, 2003. ISSN 0278-0062.
- [48] J.P. Thirion. Image matching as a diffusion process: an analogy with Maxwell’s demons. *Medical image analysis*, 2(3):243–260, 1998. ISSN 1361-8415.
- [49] J. Liu, BC Vemuri, and JL Marroquin. Local frequency representations for robust multi-modal image registration. *Medical Imaging, IEEE Transactions on*, 21(5):462–469, 2002. ISSN 0278-0062.
- [50] E. Haber and J. Modersitzki. Beyond mutual information: A simple and robust alternative. *Bildverarbeitung fr die Medizin 2005*, pages 350–354, 2005.
- [51] J. Modersitzki. *Numerical methods for image registration*. Oxford University Press, USA, 2004. ISBN 0198528418.
- [52] S. Klein, M. Staring, K. Murphy, M.A. Viergever, and J.P.W. Pluim. Elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging*, 29(1):196–205, 2010.
- [53] G. Subsol. Crest lines for curve-based warping. *Brain warping*, pages 241–262, 1999.
- [54] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, page 465. Published by the IEEE Computer Society, 1996.
- [55] S.R. Aylward, J. Jomier, S. Weeks, and E. Bullitt. Registration and analysis of vascular images. *International Journal of Computer Vision*, 55(2):123–138, 2003. ISSN 0920-5691.
- [56] S.R. Aylward and E. Bullitt. Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *Medical Imaging, IEEE Transactions on*, 21(2):61–75, 2002. ISSN 0278-0062.
- [57] G. Subsol, J.P. Thirion, and N. Ayache. Application of an automatically built 3D morphometric brain atlas: Study of cerebral ventricle shape. In *Visualization in biomedical computing*, pages 373–382. Springer, 1996.
- [58] PM Nielsen, IJ Le Grice, BH Smaill, and PJ Hunter. Mathematical model of geometry and fibrous structure of the heart. *American Journal of Physiology- Heart and Circulatory Physiology*, 260(4):H1365, 1991. ISSN 0363-6135.
- [59] M. Audette, K. Siddiqi, and T. Peters. Level-set surface segmentation and fast cortical range image tracking for computing intrasurgical deformations. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI99*, pages 788–797. Springer, 1999.

- [60] S. Sandor and R. Leahy. Surface-based labeling of cortical anatomy using a deformable atlas. *Medical Imaging, IEEE Transactions on*, 16(1):41–54, 2002. ISSN 0278-0062.
- [61] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004. ISSN 0920-5691.
- [62] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50. Manchester, UK, 1988.
- [63] A. Myronenko and X. Song. Point set registration: coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.
- [64] M.A. Audette, F.P. Ferrie, and T.M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201–217, 2000. ISSN 1361-8415.
- [65] D.G. Lowe. Object recognition from local scale-invariant features. In *International Conference on Computer Vision, ICCV*, page 1150. Published by the IEEE Computer Society, 1999.
- [66] D. Xu, H. Li, and Z. Lin. Content-based 3-D shape retrieval for pervasive computing. In *2006 1st International Symposium on Pervasive Computing and Applications*, pages 206–211, 2006.
- [67] L.S. Hibbard and R.A. Hawkins. Objective image alignment for three-dimensional reconstruction of digital autoradiograms. *Journal of neuroscience methods*, 26(1):55–74, 1988. ISSN 0165-0270.
- [68] G. Stockman. Object recognition and localization via pose clustering. *Computer Vision, Graphics, and Image Processing*, 40(3):361–387, 1987. ISSN 0734-189X.
- [69] D.P. Huttenlocher and W.J. Rucklidge. A multi-resolution technique for comparing images using the Hausdorff distance. 1992.
- [70] Y. Lamdan, J.T. Schwartz, and H.J. Wolfson. Object recognition by affine invariant matching. In *Computer Vision and Pattern Recognition, 1988. Proceedings CVPR’88., Computer Society Conference on*, pages 335–344. IEEE, 2002. ISBN 0818608625.
- [71] S. Sclaroff and A.P. Pentland. Modal matching for correspondence and recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(6):545–561, 2002. ISSN 0162-8828.
- [72] A.D.J. Cross and E.R. Hancock. Graph matching with a dual-step EM algorithm. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(11):1236–1253, 2002. ISSN 0162-8828.

- [73] A.W. Fitzgibbon. Robust registration of 2D and 3D point sets. *Image and Vision Computing*, 21(13-14):1145–1153, 2003. ISSN 0262-8856.
- [74] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *3dim*, page 145. Published by the IEEE Computer Society, 2001.
- [75] D. Demirdjian. Combining geometric-and view-based approaches for articulated pose estimation. *European Conference on Computer Vision-ECCV 2004*, pages 183–194, 2004.
- [76] L. Mundermann, S. Corazza, and TP Andriacchi. Accurately measuring human movement using articulated ICP with soft-joint constraints and a repository of articulated models. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–6. IEEE, 2007. ISBN 1424411807.
- [77] G.C. Sharp, S.W. Lee, and D.K. Wehe. Maximum-likelihood registration of range images with missing data. *IEEE transactions on pattern analysis and machine intelligence*, pages 120–130, 2007. ISSN 0162-8828.
- [78] 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. ISSN 0262-8856.
- [79] Y. Liu. Automatic 3D free form shape matching using the graduated assignment algorithm. *Pattern Recognition*, 38(10):1615–1631, 2005. ISSN 0031-3203.
- [80] Y. Liu. A mean field annealing approach to accurate free form shape matching. *Pattern Recognition*, 40(9):2418–2436, 2007. ISSN 0031-3203.
- [81] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003. ISSN 1077-3142.
- [82] A. Rangarajan, H. Chui, and F. Bookstein. The softassign procrustes matching algorithm. In *Information Processing in Medical Imaging*, pages 29–42. Springer, 1997.
- [83] A. Myronenko, X. Song, and M.A. methodarreira Perpinn. Non-rigid point set registration: Coherent Point Drift. *Advances in Neural Information Processing Systems*, 19: 1009, 2007. ISSN 1049-5258.
- [84] M. Sofka, G. Yang, and C.V. Stewart. Simultaneous covariance driven correspondence (cdc) and transformation estimation in the expectation maximization framework. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007. ISBN 1424411807.
- [85] B. Jian and B.C. Vemuri. A robust algorithm for point set registration using mixture of gaussians. 2005. ISSN 1550-5499.

- [86] H. Chui and A. Rangarajan. A feature registration framework using mixture models. In *Mathematical Methods in Biomedical Image Analysis, 2000. Proceedings. IEEE Workshop on*, pages 190–197. IEEE, 2002. ISBN 0769507379.
- [87] S. Granger and X. Pennec. Multi-scale EM-ICP: A fast and robust approach for surface registration. *European Conference on Computer Vision ECCV 2002*, pages 69–73, 2006.
- [88] W.M. Wells III. Statistical approaches to feature-based object recognition. *International Journal of Computer Vision*, 21(1):63–98, 1997. ISSN 0920-5691.
- [89] 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. ISSN 0031-3203.
- [90] D. Zhang and G. Lu. Review of shape representation and description techniques. *Pattern recognition*, 37(1):1–19, 2004.
- [91] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1615–1630, 2005.
- [92] X. Bai, Q. Li, L.J. Latecki, W. Liu, and Z. Tu. Shape band: A deformable object detection approach. 2009.
- [93] A.C. Berg and J. Malik. Geometric blur for template matching. 2001.
- [94] A. Toshev, J. Shi, and K. Daniilidis. Image matching via saliency region correspondences. In *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR'07*, pages 1–8, 2007.
- [95] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of cognitive neuroscience*, 3(1):71–86, 1991.
- [96] P. Wunsch and A.F. Laine. Wavelet descriptors for multiresolution recognition of hand-printed characters. *Pattern Recognition*, 28(8):1237–1249, 1995. ISSN 0031-3203.
- [97] X. Fu, Y. Li, R. Harrison, and S. Belkasim. Content-based image retrieval using gabor-zernike features. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 2, pages 417–420. IEEE, 2006. ISBN 0769525210.
- [98] R.J. Prokop and A.P. Reeves. A survey of moment-based techniques for unoccluded object representation and recognition. *CVGIP: Graphical Models and Image Processing*, 54(5):438–460, 1992.
- [99] D. Zhang and G. Lu. A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval. *Journal of Visual Communication and Image Representation*, 14(1):39–57, 2003.

- [100] H. Ling and D.W. Jacobs. Using the inner-distance for classification of articulated shapes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005*, pages 719–726, 2005.
- [101] P.F. Felzenszwalb and J.D. Schwartz. Hierarchical matching of deformable shapes. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007.
- [102] C. Xu, J. Liu, and X. Tang. 2d shape matching by contour flexibility. *IEEE transactions on pattern analysis and machine intelligence*, 31(1):180–186, 2009.
- [103] Z. Tu and A. Yuille. Shape Matching and Recognition—Using Generative Models and Informative Features. *European Conference on Computer Vision-ECCV 2004*, pages 195–209, 2004.
- [104] L.J. Latecki, R. Lakamper, and U. Eckhardt. Shape descriptors for non-rigid shapes with a single closed contour. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1. Citeseer, 2000.
- [105] Jamie Shotton, Andrew Blake, and Roberto Cipolla. Multiscale categorical object recognition using contour fragments. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(7):1270–1281, 2008. ISSN 0162-8828.
- [106] R.C. Veltkamp. Shape matching: Similarity measures and algorithms. In *smi*, page 0188. Published by the IEEE Computer Society, 2001.
- [107] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Transactions on pattern analysis and machine intelligence*, 10(6):849–865, 1988.
- [108] M. Schultz and T. Joachims. Learning a distance metric from relative comparisons. *Advances in Neural Information Processing Systems 16, Proceedings of the 2003 Conference*, 2003.
- [109] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.
- [110] C. Chin-Sheng, Y. Chun-Wei, and Y. Peng-Yeng. A novel Fourier descriptor based image alignment algorithm for automatic optical inspection. *J. Vis. Commun. Image R.*, 2009.
- [111] I. Kunttu, L. Lepisto, J. Rauhamaa, and A. Visa. Multiscale Fourier descriptors for defect image retrieval. *Pattern Recognition Letters*, 27(2):123–132, 2006.
- [112] I. Biederman. Recognition-by-components: A theory of human image understanding. *Psychological review*, 94(2):115–147, 1987.

- [113] T. Lindeberg. Scale-space: A framework for handling image structures at multiple scales. *CERN European Organization for Nuclear Research*, pages 27–38, 1996.
- [114] S. Sclaroff. Deformable prototypes for encoding shape categories in image databases. *Pattern Recognition*, 30(4):627–641, 1997.
- [115] G. Taubin and D.B. Cooper. Object recognition based on moment (or algebraic) invariants. *Geometric invariance in computer vision*, pages 375–397, 1992.
- [116] D. Xu and H. Li. Geometric moment invariants. *Pattern recognition*, 41(1):240–249, 2008.
- [117] V. Štruc and N. Pavešić. A comparison of feature normalization techniques for PCA-based palmprint recognition. In *Proceedings of the international conference MATHMOD*, number i, pages 2450–2453, 2009.
- [118] D. Sharvit, J. Chan, H. Tek, and B.B. Kimia. Symmetry-Based Indexing of Image Databases. *Journal of Visual Communication and Image Representation*, 9(4):366–380, 1998.
- [119] A.M. Bronstein, M.M. Bronstein, A.M. Bruckstein, and R. Kimmel. Analysis of two-dimensional non-rigid shapes. *International Journal of Computer Vision*, 78(1):67–88, 2008.
- [120] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [121] T.B. Sebastian, P.N. Klein, and B.B. Kimia. Recognition of shapes by editing their shock graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 550–571, 2004.
- [122] W. Zhang, P. Dosch, and S. Tabbone. Performance evaluation on several recognition algorithms. In *8th IAPR Workshop on Graphics Recognition*, 2007.
- [123] LeCun, Y. and Corinna, C. <http://yann.lecun.com/exdb/mnist/>. Technical report.
- [124] D. Keysers, T. Deselaers, C. Gollan, and H. Ney. Deformation models for image recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1422–1435, 2007.
- [125] D.C. Cireşan, U. Meier, L.M. Gambardella, J. Schmidhuber, R.G. Ahanger, M. Yahyazadehfar, H. Pournaghshband, S. Abraham, S. Sanyal, M. Sanglikar, et al. Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition. *Arxiv preprint arXiv:1003.0358*, 2010.

- [126] V. Walimbe and R. Shekhar. Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations. *Medical image analysis*, 10(6):899–914, 2006. ISSN 1361-8415.
- [127] X. Peng, W. Chen, and Q. Ma. Feature-based nonrigid image registration using a Hausdorff distance matching measure. *Optical Engineering*, 46:057201, 2007.
- [128] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Comparison of edge-based and ridge-based registration of CT and MR brain images. *Medical image analysis*, 1(2):151–161, 1996. ISSN 1361-8415.
- [129] M.P. Do Carmo and M.P. Do Carmo. *Differential geometry of curves and surfaces*, volume 3. Prentice-Hall Englewood Cliffs, NJ, 1976.
- [130] J.J. Koenderink. Solid shape. 1990.
- [131] J.P. Thirion and A. Gourdon. *The 3D marching lines algorithm and its application to crest lines extraction*. Citeseer, 1992.
- [132] C. Kirbas and F.K.H. Quek. Vessel extraction techniques and algorithms: A survey. In *Bioinformatics and Bioengineering, 2003. Proceedings. Third IEEE Symposium on*, pages 238–245. IEEE, 2003. ISBN 0769519075.
- [133] GP Penney, JM Blackall, MS Hamady, T. Sabharwal, A. Adam, and DJ Hawkes. Registration of freehand 3D ultrasound and magnetic resonance liver images. *Medical Image Analysis*, 8(1):81–91, 2004. ISSN 1361-8415.
- [134] YT Zhou and R. Chellappa. Computation of optical flow using a neural network. In *Neural Networks, 1988., IEEE International Conference on*, pages 71–78. IEEE, 2002.
- [135] J.C. Lee, B.J. Sheu, W.C. Fang, and R. Chellappa. VLSI neuroprocessors for video motion detection. *Neural Networks, IEEE Transactions on*, 4(2):178–191, 2002. ISSN 1045-9227.
- [136] V. Prinet, O. Monga, C. Ge, SL Xie, and SD Ma. Thin network extraction in 3d images: Application to medical angiograms. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 3, pages 386–390. IEEE, 2002. ISBN 081867282X.
- [137] D. Paquin, D. Levy, E. Schreibmann, and L. Xing. Multiscale image registration. *Mathematical biosciences and engineering*, 3(2):389, 2006. ISSN 1547-1063.
- [138] P.A. Van den Elsen, E.J.D. Pol, and M.A. Viergever. Medical image matching-a review with classification. *Engineering in Medicine and Biology Magazine, IEEE*, 12(1):26–39, 2002. ISSN 0739-5175.

- [139] JF Krucker, G.L. LeCarpentier, J.B. Fowlkes, and P.L. Carson. Rapid elastic image registration for 3-D ultrasound. *Medical Imaging, IEEE Transactions on*, 21(11):1384–1394, 2003. ISSN 0278-0062.
- [140] F. Lekien and J. Marsden. Tricubic interpolation in three dimensions. *International Journal for Numerical Methods in Engineering*, 63(3):455–471, 2005. ISSN 1097-0207.
- [141] E.B. van de Kraats, G.P. Penney, D. Tomazevic, T. van Walsum, and W.J. Niessen. Standardized evaluation methodology for 2-D-3-D registration. *IEEE Transactions on Medical Imaging*, 24(9):1177, 2005. ISSN 0278-0062.
- [142] B.K.P. Horn, H.M. Hilden, and S. Negahdaripour. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America A*, 5(7):1127–1135, 1988.
- [143] OD Faugeras and M. Hebert. The representation, recognition, and locating of 3-D objects. *The international journal of robotics research*, 5(3):27, 1986.