A Technique for Face Recognition Based on Image Registration

by

Steven Gillan
B.Eng, University of Victoria, 2008

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of

M A S T E R O F A P P L I E D S C I E N C E

in the Department of Electrical and Computer Engineering

© Steven Gillan, 2010
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part by photocopy or other means, without the permission of the author.
A Technique for Face Recognition Using Image Registration

by

Steven Gillan
B.Sc., University of Victoria, 2008

Supervisory Committee

Dr. P. Agathoklis, Supervisor
(Department of Electrical and Computer Engineering)

Dr. W.S. Lu, Departmental Member
(Department of Electrical and Computer Engineering)
Supervisory Committee

Dr. P. Agathoklis, Supervisor
(Department of Electrical and Computer Engineering)

Dr. W. S. Lu, Departmental Member
(Department of Electrical and Computer Engineering)

ABSTRACT

This thesis presents a technique for face recognition that is based on image registration. The image registration technique is based on finding a set of feature points in the two images and using these feature points for registration. This is done in four steps. In the first, images are filtered with the Mexican hat wavelet to obtain the feature point locations. In the second, the Zernike moments of neighbourhoods around the feature points are calculated and compared in the third step to establish correspondence between feature points in the two images and in the fourth the transformation parameters between images are obtained using an iterative least squares technique. The face recognition technique consists of three parts, a training part, an image registration part and a post-processing part. During training a set of images are chosen as the training images and the Zernike moments for the feature points of the training images are obtained and stored. In the registration part, the transformation parameters to register the training images with the images under consideration are obtained. In the post-processing, these transformation parameters are used to determine whether a valid match is found or not.

The performance of the proposed method is evaluated using various face databases and it is compared with the performance of existing techniques. Results indicate that the proposed technique gives excellent results for face recognition in conditions of varying pose, illumination, background and scale. These results are comparable to other well known face recognition techniques.
Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents iv
List of Tables vii
List of Figures viii
List of Abbreviations x
Acknowledgement xi
Dedication xii

1 Introduction 1
1.1 Image Registration .................................................. 1
1.2 Face Recognition ....................................................... 3
1.3 Comparison of Various Face Recognition Techniques ................. 4
1.4 Scope and Contributions of the Thesis ................................ 5

2 Image Registration 7
2.1 Abstract .................................................................. 7
2.2 Introduction ............................................................. 7
2.3 Feature Point Extraction ............................................. 9
# Table of Contents

2.3.1 Effect of Image Scale on Feature Extraction ....................................... 12

2.4 Determining Correspondence Between Images ................................. 18
  2.4.1 Zernike Moment Calculation of Feature Points ........................ 19
  2.4.2 Use of Correspondence Matrix ........................................ 20

2.5 Transformation Parameters Estimation ........................................... 21
  2.5.1 Iterative Weighted Least-Squares Algorithm .......................... 23

2.6 Examples .................................................................................... 26

2.7 Summary ..................................................................................... 35

3 Face Recognition ............................................................................. 37
  3.1 Abstract .................................................................................... 37
  3.2 Introduction ................................................................................ 37
  3.3 Application of Image Registration for Face Recognition .......... 38
  3.4 Mexican Hat Based Algorithm for Face Recognition ............... 40
    3.4.1 Training Step .................................................................. 40
    3.4.2 Post-Processing .............................................................. 42
  3.5 Demonstration of the Proposed Method ..................................... 46
    3.5.1 Demonstration Against Cases of Varying Pose ................. 46
    3.5.2 Demonstration Against Cases of Varying Illumination .... 48
    3.5.3 Demonstration Against Cases of Varying Background and Scale . 48
  3.6 Summary .................................................................................... 52

4 Experimental Design and Results ..................................................... 54
  4.1 Abstract .................................................................................... 54
  4.2 Introduction ................................................................................ 54
  4.3 Experiments using the Yale Face Database B ......................... 56
    4.3.1 Experiment 1: Effect of Varying Illumination with Constant Pose 59
    4.3.2 Experiment 2: Varying Pose and Constant Illumination Condition . 60
    4.3.3 Experiment 3: Effect of Varying Illumination and Varying Pose .. 62
4.3.4 Comparison of Results With Other Algorithms .......................... 63
4.4 Experiment using the Caltech Face Database .............................. 65
  4.4.1 Caltech Face Database .................................................. 66
  4.4.2 Performance of the Proposed Method on the Caltech Database .. 67
4.5 Summary .................................................................................. 69

5 Conclusion and Future Work ...................................................... 70
  5.1 Conclusions ............................................................................ 70
    5.1.1 Image Registration ......................................................... 70
    5.1.2 Face Recognition ........................................................... 71
    5.1.3 Experimental Results ..................................................... 72
  5.2 Future Work ........................................................................... 73

Bibliography ................................................................................. 75
# List of Tables

<table>
<thead>
<tr>
<th>Table Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of Non-Scale Image Registration Techniques</td>
<td>35</td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison of Scale based Image Registration Techniques</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Average comparison time in seconds per image</td>
<td>58</td>
</tr>
<tr>
<td>4.2</td>
<td>Total Error summary for Experiment 1</td>
<td>60</td>
</tr>
<tr>
<td>4.3</td>
<td>Breakdown of the False Acceptance and False Rejection rate for Experiment 1</td>
<td>60</td>
</tr>
<tr>
<td>4.4</td>
<td>Total Error summary for Experiment 2</td>
<td>61</td>
</tr>
<tr>
<td>4.5</td>
<td>Breakdown of the False Acceptance and False Rejection rate for Experiment 2</td>
<td>61</td>
</tr>
<tr>
<td>4.6</td>
<td>Comparison of Results for Experiment 2</td>
<td>62</td>
</tr>
<tr>
<td>4.7</td>
<td>Total Error summary for Experiment 3</td>
<td>62</td>
</tr>
<tr>
<td>4.8</td>
<td>Breakdown of the False Acceptance and False Rejection rate for Experiment 3</td>
<td>63</td>
</tr>
<tr>
<td>4.9</td>
<td>Comparison of Results for Experiment 1</td>
<td>64</td>
</tr>
<tr>
<td>4.10</td>
<td>Comparison of Results for Experiment 3</td>
<td>65</td>
</tr>
<tr>
<td>4.11</td>
<td>Results from Caltech Database</td>
<td>68</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Sample images showing partial overlap ......................................... 8
2.2 Summary of the Image Registration Algorithm. ................................. 9
2.3 1-D Mexican-Hat Wavelet .......................................................... 10
2.4 2-D Mexican Hat Wavelet .......................................................... 10
2.5 An example of the feature extraction step. ...................................... 13
2.6 Feature Point Extraction with Scaling Example ............................... 14
2.7 Effect of Scaling Parameters ....................................................... 16
2.8 Scale ratio distribution ............................................................. 17
2.9 Zernike polynomials ................................................................. 18
2.10 Transformation example ............................................................ 24
2.11 Example: Ikonos acquired images ................................................ 28
2.12 Example: Landsat Thematic Mapper (TM) Bands ............................ 29
2.13 Example: Mountain Landscape ................................................... 31
2.14 Example: Paris .................................................................. 32
2.15 Example: Noise effected images ................................................... 34

3.1 Summary of the Face Recognition Algorithm ................................. 39
3.2 Training step procedure ............................................................. 41
3.3 Error Parameter vs. Scale Parameter ............................................. 45
3.4 Face Recognition Demonstration versus Pose ................................ 47
3.5 Face Recognition Demonstration versus Illumination ...................... 49
3.6 Face Recognition Demonstration versus Background and Scale ........ 50

4.1 Yale Face Database B ................................................................. 56
List of Figures

4.2 Sample Images of Caltech Database .......................... 66
4.3 Example of Rejected Images from Caltech Database ............. 67
List of Abbreviations

1-D       One Dimension
2-D       Two Dimension
CCTV      Closed Circuit Television
CT        Computed Tomography
FFT       Fast Fourier Transform
ICA       Independent Component Analysis
LoG       Laplacian of Gaussian
MRI       Magnetic Resonance Imaging
PCA       Principal Component Analysis
PET       Positron Emission Tomography
RANSAC    Random Sample Consensus
RMSE      Root Mean Squared Error
SIFT      Scale Invariant Feature Transform
SVD       Singular Value Decomposition
TM        Thematic Mapper
Acknowledgement

I would like to thank:

Dr. Pan Agathoklis, for mentoring, support, encouragement, and patience.

my fellow graduate peers, Adam and Fil, for helping to keep me sane throughout all of this.

University of Victoria, Department of Electrical and Computer Engineering, for funding me with a University Fellowship and providing me with teaching opportunities.

Whatever course you decide upon, there is always someone to tell you that you are wrong. There are always difficulties arising which tempt you to believe that your critics are right. To map out a course of action and follow it to an end requires courage.

Ralph Waldo Emerson
Dedication

To my wife Roslyn, your support and belief was instrumental to me. To my Mom, her strength and courage is an inspiration.
Chapter 1

Introduction

1.1 Image Registration

Image registration is the process of estimating the parameters of the geometric transformation model that can be used to map a given (target) image to an original (reference) image. An overview of image registration techniques are found in [1–3]. In general, image registration techniques fall within two methodologies: area based and feature based.

Area based methods estimate the transformation between two images by analyzing pixel intensities of an image using various properties. These properties can include mutual information of the images. Mutual information is a concept from information theory which is the measure of dependence between two images based on the assumption that the gray values are of maximum dependence when the images are correctly aligned. Mutual information is a popular method for medical image registration due to the generality of the algorithm. In [4] mutual information is used to register various medical imagery (CT, PET, and MRI) of brain scans and also rigid body scans. The technique from [5] uses mutual information and template matching to improve on the computation time by comparing the mutual information between subimages of a target image to a template of a reference image rather than the entire image. In [6], an updated method is used that makes use of Parzen windows and a gauss kernel function to determine the mutual information. The computation time is greatly reduced by using the Fast Gauss Transform. The technique in [7], illustrates the use of mutual information for remote sensing application using a joint
histogram estimation algorithm and B-splines as the kernel function.

Other area based methods come from the Fourier transform. In [8], an extension of the phase correlation technique is applied which matches images with differences in translation, rotation and scale within the Fourier domain. While FFT based methods are difficult for cases where scale changes between the images exist, the method in [9] uses a Pseudo-Polar Fourier Transform and improves on the amount of scaling that can be achieved while also greatly reducing the computation time of the algorithm. A similar method is presented in [10] which uses the Pseudo-Log-Polar Fourier Transform to estimate rotation, scaling and translation effects.

Feature based methods use the extraction of feature points of an image to provide a means of correlation of areas between the two images which are similar and are then used to obtain the parameters for transformation. The methods based on this technique use image characteristics such as edges, corners, feature points, line segments, contours and curvature of image intensity. The technique from [11] presents a registration technique using closed-boundary regions for the feature point extraction. Based on finding the contours of an image and calculating the center of gravity of these areas to be used as the feature points. In [12], an automated matching based on cross correlation is used to find the control points of an image. Then using a robust estimation of these control points provided by the random sample consensus algorithm (RANSAC). Another technique which uses edge-based selections to determine the control points of an image is presented in [13, 14]. Correspondence using this method is found using template matching while the transformation between the two images is performed using spline interpolation. Further improvements to determine the correspondence of the feature points between two images using constraints in both spatial relations and feature similarity [15]. A method which is popular in computer-vision applications is the scale invariant feature transform (SIFT) which is based on the feature point detection in scale-space. This technique has shown to be very robust for difference in scale and affine distortions applied to the images.

Recently an interesting technique used in the application of image registration has been
developed which is based on the scale interaction of Mexican-hat wavelets [16, 17]. Feature points extracted from the images are then used to provide a correspondence between an original (reference) image and a second (target) image by comparing the magnitudes of the Zernike moment calculated around these feature points [18, 19]. These same feature points will also be used to provide the normalization parameters required to transform the images to a standard form. This technique can be applied in many applications such as medical imaging, remote sensing, vision, and photography. In this thesis, this method will be applied to face recognition.

1.2 Face Recognition

In the last couple years, the importance of face recognition algorithms has increased [20, 21]. In the realm of biometric identification that currently exist today, face recognition offers a unique model as it does not require the subject to have first-hand knowledge that is it being applied. Increasing applications of closed circuit television (CCTV) and other forms of surveillance add a premium on being able to effectively identify and validate persons within a clip. These applications can also require added demands of trying to identify further features of the subjects, gender, ethnicity, age, etc. With these considerations, it becomes important that when developing a face recognition algorithm there is a level of robustness available such that variance in the lighting conditions and physical interactions of the subjects does not severely affect the performance of the algorithm [22]. Furthermore, changes in facial features (glasses, facial hair, expression) are also important considerations to the success of the face recognition algorithm [20].

Currently, there are a number of interesting techniques that are used in face recognition. Specifically, these can be divided into two types of methods; feature-based techniques and appearance-based techniques. From the methods that are appearance based, Principal Component Analysis (PCA) and its variant, known as Eigenfaces [23–26], has been very popular and has been shown to be very effective. The Eigenfaces, simply put, is a method
to represent a face as a linear combination of basis images, then using these principal components, any face in the set can be recreated with a high level of accuracy. This method has many advantages. Among them is the speed of the algorithm since by representing an image in multiple lower dimensional spaces, the classification method can work much faster. A simpler method that can be performed is a direct nearest neighbor matching within the image space, which when used with images that have been normalized with zero mean and unit variance is known as the Correlation method[27]. However, these methods are prone to very high error when the subject image has changes in lighting illumination and pose. An illumination cone representation was presented [28, 29] which showed very good results with respect to variation in both pose and illumination conditions. This method uses as little as three training images of a fixed pose and different but unknown illumination and creates a 3-D model of the face.

Feature based techniques use the interaction of feature points extracted from an image to determine whether correspondence between two images is present. Some of these techniques use feature extraction using Gabor wavelets [30–34]. These Gabor wavelets have presented some good results in robustness to differences in facial expression and pose. Another method of feature based face recognition is based on using Scale Invariant Feature Transform (SIFT) [35–39] and was adapted from the broader application of object recognition using SIFT[40, 41]. In [35, 36] it was shown that face recognition methods that are based on the SIFT technique of image registration allows for matching across a large range of illuminations and affine distortions.

1.3 Comparison of Various Face Recognition Techniques

To determine the effectiveness of the developed face recognition method and compare it with other methods, data generated from a set of experiments are being used. The means to perform these experiments are generated using a database of images. These experiments are designed to evaluate the various face recognition techniques to conditions such as dif-
ferences in pose, illumination, scale, location, gesture, etc. The results of the experiments using a specific database of images provides a measure of how well the face recognition method compares to other existing techniques. The first type of example for evaluation the face recognition methods is by using a controlled set of data where the faces are consistent throughout all the images. This evaluation measures the face recognition method strictly against varying conditions of pose and illumination. This example makes use of the Yale Face Database B [29] which was specifically designed for evaluating the performance of face recognition algorithms specifically to changes in pose and illumination conditions. Another example which will be used will be in an uncontrolled setting, where the database images are not controlled with respect to background, scale, position, expression, illumination and pose. This example will make use of the Caltech face database [42, 43] which contain images that are acquired in an uncontrolled environment.

1.4 Scope and Contributions of the Thesis

This thesis is organized as follows:

In Chapter 2, the image registration technique proposed in [16, 17, 44] is presented. This technique has been further improved with respect to the implementation to increase the efficiency of the algorithm. The choice of scale parameters for the feature point extraction step has also been modified to use a smaller number of filtering operations while maintaining a large number of scale comparisons. This image registration technique is based on a feature extractor using scale interactions of Mexican-hat wavelets to find feature points on an image. This chapter will also discuss the effect that the scale of an image has on the feature extraction. The magnitude of the Zernike moments of these feature points are then used to determine correspondence with other images. Once corresponding feature point pairs are found, an iterative weighted-least squares minimization is performed to determine the transformation parameters to register the two images and remove any outlier feature point pairs that may exist.
1.4 Scope and Contributions of the Thesis

In Chapter 3, a face recognition method based on the image registration technique is developed. This face recognition method consists of a training step, the feature extraction and registration step (using the image registration technique of chapter 2) and a post-processing step. The training step is performed to help reduce the effects that pose and illumination have on the face recognition problem, and also increase the speed of the face recognition technique. The post-processing step is used to determine a correct match or rejection between the images being compared.

In Chapter 4, the performance of the face recognition technique presented in chapter 3 is evaluated using face image databases. These image databases are designed to test against changes in pose, illumination, background and scale. Two sets of image databases are used, the Yale Face Database B [29] consists of images of varying pose and illumination. The Caltech Face Database [42, 43] consists of images of varying background, pose, posture, scale and facial expression. The results of the experiments conducted using these image databases are then compared to the performance of other existing face recognition techniques. The results presented in this chapter show that the image registration technique provides a good approach to face recognition.

Finally, in Chapter 5, the results and contributions of this thesis are summarized and directions for future research areas on this topic are suggested.
Chapter 2

Image Registration

2.1 Abstract

This chapter will introduce the image registration technique to be used in the next chapter for face recognition. Specifically, a feature-based technique which uses Mexican-hat wavelets to determine feature points in an image. The chapter is organized as follows: The feature point extraction step including the use of scale interactions of Mexican-hat wavelets and considerations to the effect of image scale effects are discussed in section 2.3. Section 2.4 introduces the use of the magnitudes of Zernike moments of the feature points to determine correspondence between the feature points of two images. Section 2.5 presents an iterative weighted least-squares minimization to eliminate any outlier pairs and to find the transformation parameters needed to join the two images. Section 2.6 presents a number of different examples illustrating how the image registration technique being used in different conditions.

2.2 Introduction

Image registration is the process of estimating the parameters of the geometric transformation model that can be used to map a given (target) image to an original (reference) image [1, 2]. In general, image registration techniques fall within two methodologies: area based and feature based. Area based methods estimate the transformation between two
images by analyzing pixel intensities of an image using various properties. Feature based methods use the extraction of feature points of an image to provide a means of correlation between the feature points of the two images which are then used to obtain the parameters for transformation.

![First Aerial Test Image](image1.png) ![Second Aerial Test Image](image2.png)

**Figure 2.1:** Sample images showing partial overlap

An example of image registration is shown in the application of aerial photography where aerial images from an airplane or satellite are used to map coverage of an area. Due to different directions that the images can be acquired from, partial overlap may occur, or due to different weather conditions taken between images, it can be difficult to correctly match the images together. Figure 2.1a and Figure 2.1b show one such scenario where partial overlap occurs and the images have been taken at different orientations. These images will be used as two of the test images to illustrate the image registration algorithm.

In the remainder of this chapter a detailed description of each of the elements of the image registration algorithm, which is shown graphically in Figure 2.2, is outlined. This proposed technique was developed in [16, 17, 44] and was adapted here with improvements to the choice of image scale parameters and efficient, fast implementation using C.
2.3 Feature Point Extraction

Figure 2.2: Summary of the Image Registration Algorithm.

The algorithm itself is explained using 4 steps. First the feature points of the images to be registered are found, using scale interactions Mexican-hat wavelet for filtering the image while allowing for compensation in scale differences between images. Next, the Zernike moments are determined within circular neighborhoods around each of the features points which are then used for determining the correspondence between the images. This correspondence is based on the similarity of the magnitudes of the Zernike moments between the two images. Finally, an iterative weighted least-squares minimization is performed in order to remove any outlier points, as well as to determine a set of affine parameters that can be used to transform one image onto the other. Some examples of the image registration algorithm being used on different sample images will be shown and discussed at the end of this chapter.

2.3 Feature Point Extraction

Feature points are locations in an image that represent a local maxima of a certain function within a neighborhood of image points on an image. Used in an image registration application, these feature points become the locations in an image which can be used to determine whether two images have matching points and also in estimating the transformation parameters needed to combine two images.

There are many approaches to find feature points. The method presented in this thesis uses a two-step process [44, 45].

1. Comparison of the response of an image with a Mexican-hat wavelet applied to it.
2. Searching for local maxima within the response.
2.3 Feature Point Extraction

Figure 2.3: 1-D Mexican-Hat Wavelet

(a) $s_m = 1$
(b) $s_m = 2$
(c) $s_m = 4$

Figure 2.4: 2-D Mexican-hat Wavelet. The top row is the frequency response of the wavelet, the bottom row is the space-domain response of the wavelet at different values of $s_m$
A Mexican-hat wavelet, shown by its 1-D response in Figure 2.3, is the negative second derivative of a Gaussian function, also called the Laplacian of a Gaussian (LoG). The name comes from the shape of the wavelet, which has a sharp positive peak surrounded by a negative trough, which resembles that of a traditional Mexican hat. For the purpose of image registration, a 2-D version is used, which is shown in Figure 2.4.

Applying a Mexican-Hat wavelet to an image will serve to pronounce some areas of an image, while at the same time reducing the intensity of the values around it. This helps in accurately identifying the exact location of the features on an image. The response of the image can be represented mathematically [44]:

\[
\phi(x, s_1, s_2) = |\Re(x, s_1) - \Re(x, s_2)|
\]  

(2.1)

where \(x = (x_1, x_2)\) represents the vertical and horizontal coordinates of the image and

\[
\Re(x, s_m) = I(x) \otimes \text{Mex}(x, s_m)
\]

which is simply the image, \(I(x)\), convolved with the Mexican-hat wavelet which is represented

\[
\text{Mex}(x, s_m) = \frac{1}{2^{-s_m}} \left( 2 - \frac{x_1^2 + x_2^2}{2^{-s_m}} \right) e^{-\frac{1}{2} \frac{x_1^2 + x_2^2}{2^{-s_m}}}
\]  

(2.2)

where \(s_m\) represents the scaling value used in the function.

Performing the two-dimensional (2-D) convolution using the C programming language may not be very efficient. Therefore, applying the property of space convolution [46], convolution in the space domain can be performed as multiplication in the frequency domain which can be easily programmed using signal processing libraries[47, 48].

\[
f(t_1, t_2) \otimes g(t_1, t_2) \leftrightarrow F(j\omega_1, j\omega_2)G(j\omega_1, j\omega_2)
\]

The result of applying the Mexican-hat wavelet is shown in Figure 2.5, where the different scale interactions, \(s_m\), of the Mexican-hat wavelet has on the response is seen in Figure 2.5b and Figure 2.5c. The overall response from (2.1) is shown in Figure 2.5d, which is the absolute difference between Figure 2.5b and Figure 2.5c.
Using this response, the maxima within a neighborhood can be detected and used as a feature point. This is done using a two step process:

1. Divide the image into an equal number of blocks. Search these blocks for the maximum value which will represent the possible feature points.
2. Using these maximum points as a center, search in a circular-shaped neighborhood. The maximum value within this neighborhood will be considered the strongest feature point.

The resulting maxima found from this step can be seen as the feature points that are overlaid on Figure 2.5e.

### 2.3.1 Effect of Image Scale on Feature Extraction

When dealing with two images that have a difference in scale between them, the effect that filtering with the Mexican-hat wavelet has on the two images will be different. This will result in a failure to correctly extract corresponding feature points from the two images.

In order to compensate for scaling of the pixel data in the image, the Mexican-hat wavelet from (2.2) can be redefined by associating a scaling factor $s_{pi}$.

$$
\text{Mex}(\frac{x}{s_{pi}}, s_m) = \frac{1}{2^{-s_m}} \left( 2 - \frac{(\frac{x_1}{s_{pi}})^2 + (\frac{x_2}{s_{pi}})^2}{2^{-s_m}} \right) e^{-\frac{(\frac{x_1}{s_{pi}})^2 + (\frac{x_2}{s_{pi}})^2}{2^{-s_m}}} \right) (2.3)
$$

The feature point extraction step is then applied multiple times using a range of values for $s_{pi}$ which will compensate for the difference in scale between the two images. An example of this process is shown in Figure 2.6. The feature point extraction step is performed multiple times for each of the different values of $s_{pi}$ which are shown in Figure 2.6c-2.6e. The size of the circles overlaid on these figures represent the different scale factors of $s_{pi}$ that were used to find the feature points. The overall result of the feature extraction step with scale consideration is shown in Figure 2.6b where all the feature points are shown.
2.3 Feature Point Extraction

Figure 2.5: An example of the feature extraction step.
Figure 2.6: An example of the feature point extraction step performed at different scaling values, evident in the different sized circles around the feature points. (a) The original image. (b) Final result of feature extraction with all scales. (c) $s_{p_i} = 0.8$. (d) $s_{p_i} = 1.0$. (e) $s_{p_i} = 1.25$
2.3 Feature Point Extraction

together. As smaller values of $s_{p_i}$ are used, more feature points are found on the image due to the smaller size of the local neighborhood used in the search for the local maxima.

The effect that $s_{p_i}$ has on the feature point extraction step is shown in Figure 2.7. The images on the left and right hand side are scaled down/up 25% from the middle image respectively. The top row of Figure 2.7 has the value of $s_{p_i}$ changed by the same factor. The result is that the feature points in Figure 2.7a, 2.7b and 2.7c are all the same. Applying the same value of $s_{p_i}$ on the same images, which is done in Figure 2.7d, 2.7e and 2.7f, results in inconsistent feature points being detected for each image.

The addition of the $s_{p_i}$ parameter to the feature extraction step increases the computation time of the image registration algorithm and is the main source of time spent in the algorithm. As demonstrated in Figure 2.7, the value of $s_{p_i}$ is of less importance than the ratio between the values of $s_{p_i}$ for each image. The choice of values for $s_{p_i}$ to be used for the feature point extraction step is required to ensure that an efficient and exhaustive range of scale ratios are considered. A method to create a range of $s_{p_i}$ values that are similar to a logarithmic distribution is done using conjugate powers. This range of $s_{p_i}$ values covers different scale ratios while producing a dense sampling of feature points on both images and also covers unity scaling. The resulting scaling function is shown in (2.4) below.

$$n = 2k + 1$$
$$s_{p_1} = \begin{bmatrix} a^{-k} & a^{-k+1} & \ldots & 1.0 & \ldots & a^{k-1} & a^k \end{bmatrix}$$
$$s_{p_2} = \begin{bmatrix} b^{-k} & b^{-k+1} & \ldots & 1.0 & \ldots & b^{k-1} & b^k \end{bmatrix}$$ (2.4)

where $a$ and $b$ represent the scale values to use and are generally close to each other and are typically close to 1.0, and $n$ represents the number of scale parameters in $s_{p_1}$ and $s_{p_2}$ and is always chosen to be an odd number. This method for determining the scale parameter $s_{p_i}$ will account for $n^2$ scale ratios at the cost of only performing $2n$ filtering operations. These parameters are defined by the user and can be modified to suit the scaling range of the images that are being compared. An example of the values of $s_{p_i}$ and ratios that are
Figure 2.7: The result of applying $s_{p_i}$ to scaled images. The top row adjusts $s_{p_i}$ by the same ratio that the image is scaled by, whereas the bottom row uses the same value of $s_{p_i}$ for all images.
2.3 Feature Point Extraction

Figure 2.8: An example of using $n = 5$ to produce 25 scale ratios, 12 below unity, unity scale, and 12 above unity. The step size is large for large scale ratios ($> 1$), while small for finer scale ratios ($< 1$).

obtained when using a length of $n = 3$, $a = 1.2$ and $b = 1.25$ is shown below:

$$s_{p_1} = \begin{bmatrix} 0.8333 & 1.0 & 1.2 \end{bmatrix} \quad s_{p_2} = \begin{bmatrix} 0.8 & 1.0 & 1.25 \end{bmatrix}$$

$$ratios = \frac{s_{p_1}}{s_{p_2}} = \begin{bmatrix} 0.6667 & 0.8 & 0.8333 & 0.96 & 1.0 & 1.0417 & 1.2 & 1.25 & 1.5 \end{bmatrix}$$ \hspace{1cm} (2.5)

From (2.5), it is seen that the length of each scale vector is $n$ and there are $n^2$ ratios that will be compared. Choosing $a \neq b$ ensures that there is no redundancy in scale ratios. This method for determining the scales also produces an even number of ratios above and below unity, while also adjusting the distance between these ratios according to the size of the ratio. A visual example of this is shown in Figure 2.8 and shows that a wide range of ratios are covered that focus closely around unity scaling while also testing for more extreme scale differences. Using such a method to choose the values of $s_{p_i}$ for the feature extraction step provides the maximum amount of unique ratios for comparison, while also maintaining unity scaling for cases that do not have scale interactions.

By comparison using a linear range of values for $s_{p_i}$, there exists a lot of inefficiency in performing extra iterations of the feature extraction step due to repeating the same scale ratios. Using a logarithmic scale will result in the $s_{p_i}$ values between the two images being uneven, one image having $s_{p_i}$ that are greater then unity, while the other image had scale values that were much smaller then unity. This results in an uneven sampling of feature points between the two images (one is densely samples, while the other is sparesly sampled). Logarithmic sampling also did not account well for images that had no scaling.
2.4 Determining Correspondence Between Images

Figure 2.9: An example of a 5th order Zernike polynomial. Each radial order represents more detailed responses, which when used to project an image disk will retain more information. (a) shows the form that the Zernike polynomial models in optics. (b) shows the same Zernike polynomial projected onto the unit disk.

differences in them, which affected the performance of once successful registrations.

2.4 Determining Correspondence Between Images

With the feature points for each image found as described in the previous section, the correspondence between the images can be performed. This correspondence is achieved through determining matching feature point pairs between two images. This is obtained using descriptor vectors $P_d$ defined as

$$P_d = \frac{1}{s_{p_i}^2} (|Z_{1,1}|, \ldots, |Z_{p,q}|) \quad (2.6)$$

where $|Z_{p,q}|$ is the magnitude of the Zernike moment for a feature point. A corresponding matching feature point pair between two images is found based on the minimum distance among descriptor vectors, $P_{d_i}$ and $P'_{d_j}$. 
A Zernike moment is a projection of an image disk onto an orthogonal basis of Zernike Polynomials. Zernike polynomials are used to represent waveform data in polynomial form. These polynomials share many properties of aberrations that are observed in optical imaging tests, which are shown in Figure 2.9. Projecting an image disk onto each of these different polynomials will create a linear independent Zernike moment. A collection of Zernike moments become a linear independent set of basis images that can then be used to reconstruct the image disk. A higher order of Zernike moments will maintain more detail of the original image disk. Zernike moments are used in pattern recognition and image analysis.

The advantage of using Zernike moments to determine correspondence is two fold. First, the magnitudes of the Zernike moments are rotationally invariant \[ \hat{Z}_{p,q} = Z_{p,q} e^{j\phi} \] thus the magnitudes of the Zernike moments \( Z_{p,q} \) and \( \hat{Z}_{p,q} \) are the same. Second, the Zernike moments of an image, \( I \), is related to the Zernike moments of the resized image, \( \hat{I} \) using the scaling parameter \( s_{p_i} \) \[ Z_{p,q} = \frac{1}{s_{p_i}^2} \hat{Z}_{p,q} \] (2.8)

### 2.4.1 Zernike Moment Calculation of Feature Points

Using the rotation invariance of the Zernike moment, a circular neighborhood with each feature point as the origin of the neighborhood is used as the image disk to project onto the Zernike Polynomials, each pixel in this neighborhood can be normalized to a value within the unit circle \[ \sqrt{x_1^2 + x_2^2} \leq 1. \]

\(^1\)The rotation of the images that are being compared do not need to be normalized before comparing Zernike moments.
The Zernike moment of order \( p \) is defined \([44]\) 

\[
Z_{pq} = \frac{(p + 1)}{\pi} \sum_{x_1} \sum_{x_2} V_{pq}^*(r, \theta) A(x_1, x_2)
\]

where \( A(x_1, x_2) \) is the value of the pixel intensity at the point \( x_1 \) and \( x_2 \). \( V_{pq}^* \) is the complex conjugate of the Zernike polynomial of order \( p \) and repetition \( q \). 

\[
V_{pq} = R_{pq}(r)e^{(iq\theta)}
\]

where \( r \) and \( \theta \) are the polar coordinates of \( x_1 \) and \( x_2 \)

\[
\begin{align*}
    r &= \sqrt{x_1^2 + x_2^2} \\
    \theta &= \tan^{-1}\left(\frac{x_2}{x_1}\right)
\end{align*}
\]

and \( R_{pq}(r) \) is the radial polynomial defined as

\[
R_{pq}(r) = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s(p-s)!r^{p-2s}}{s! \left(\frac{p-2s+|q|}{2}\right)! \left(\frac{p-2s-|q|}{2}\right)!}
\]

where \( p = 0, 1, 2, \ldots, \infty \), \( 0 \leq |q| \leq p \) and \( p - |q| \) is even.

High orders of Zernike moments provide information on fine details of the image, at a cost of being noise-sensitive compared to lower order Zernike moments\([18, 19, 51]\). In \([44]\), the highest order used was 10 to provide a compromise between noise sensitivity and information content of the moments.

### 2.4.2 Use of Correspondence Matrix

Once all the Zernike moments for each feature point are calculated they can be compared to the moments of each feature point on the other image and are stored in a correspondence matrix \( C^2 \). Each entry \( c_{ij} \) is the \( \ell_1 - norm \) of the difference between two descriptor vectors, \( P_d \), defined by (2.6) for two images.

\( ^2 \)Also called a distance matrix \([44]\)
2.5 Transformation Parameters Estimation

\[ c_{ij} = \ell_1(P_{d_i} - P_{d_j}) \]
\[ = \sum_{m=1}^{36} |P_{d_i}(m) - P_{d_j}(m)| \]

Where \( P_d \) is the descriptor vector of the magnitudes of the Zernike moments for feature points \( i = 1, 2, \ldots, N \) and \( j = 1, 2, \ldots, N' \).

Within this correspondence matrix, the minimum distance value along each row and column is found, shown in (2.9).

\[
\text{row}_i = \text{index} \left\{ \min_j c_{ij} \right\} \\
\text{col}_j = \text{index} \left\{ \min_i c_{ij} \right\} (2.9)
\]

A correspondence occurs when the minimum value along a row is also the minimum value on the associated column in \( C \), such that \( \text{row}_i = \text{col}_j \). In order to reduce the number of false correspondences that are found, a threshold between the lowest value, and the next lowest value is used. This results in \( K \) feature point pairs, where \( K \leq \min(N, N') \).

2.5 Transformation Parameters Estimation

The transformation parameters are used to transform one of the images to the necessary size, orientation and position to ensure that the two images are combined into one maintaining all the information individual to each image. The method to estimate these transformation parameters is performed by solving an iterative weighted least squares minimization problem. The objective function is considered to be the \( \ell_2 \) norm of the weighted errors [44].

\[
\Psi(z) = \ell_2^2(W(f(P', z) - P)) = \sum_{i=1}^{K} w_i \| f(P'_i, z) - P_i \|^2 \quad (2.10)
\]

where \( W \) represents a diagonal matrix with elements \( w_i \), the weights associated with the distance of the feature point pairs between the two images (\( P' \) and \( P \)), \( z \) are the estimated
transformation parameters. The transformation parameters are determined by solving the optimization problem:

\[
\min_z \Psi(z)
\]  

(2.11)

where \( \Psi(z) \) is the error between the feature points in the reference image and the transformed target image using the updated least-square solution.

The method for determining the weights are found iteratively, and will be discussed later in this section.

The matching transformations considered are all 2D geometric affine transformations: scaling, rotation, skewing, translation, etc. From these distortions, a set of transformation parameters can be found such that points \( x_i = (x_1, x_2) \) of image \( I \) can be mapped to the points \( \hat{x}_i = (\hat{x}_1, \hat{x}_2) \) of image \( \hat{I} \).

\[
\hat{x}_i = \begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2
\end{bmatrix} = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
t_{x1} \\
t_{x2}
\end{bmatrix}
\]  

(2.12)

Applying this transformation to a set of \( k \) feature points, then (2.12) can be set up as an overdetermined set of equations given by:

\[
\begin{bmatrix}
w_1 \\
\vdots \\
w_k
\end{bmatrix} \begin{bmatrix}
\hat{x}_1^T \\
\vdots \\
\hat{x}_k^T
\end{bmatrix} = \begin{bmatrix}
w_1 \\
\vdots \\
w_k
\end{bmatrix} \begin{bmatrix}
x_i^T 1 \\
\vdots \\
x_k^T 1
\end{bmatrix} \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix} \begin{bmatrix}
t_{x1} \\
t_{x2}
\end{bmatrix}
\]

\[
W B = W A Z
\]  

(2.13)

Where \( B \) are the coordinates of the feature points for image \( \hat{I} \), \( A \) are the coordinates for the feature points of image \( I \) and \( W \) is a weighting matrix used to distinguish between outlier and inlier feature point pairs. The elements of \( W, w_i \), will be obtained using an iterative algorithm discussed in the next section. The matrix \( Z \) can be obtained by finding the least
2.5 Transformation Parameters Estimation

squares solution to (2.13) using the singular value decomposition (SVD) [52] of $WA$.

$$WAZ = WB$$

$$UΣV^T Z = WB$$

$$Z = VΣ^{-1}U^T WB$$  \hspace{1cm} (2.14)

A similar method for solving for the affine parameters, $Z$, was performed in [45] which made use of a closed form invertible matrix which is found from solving the least-squares solution such that the gradient of the objective function is zero. While, this method allows for the use of matrix inversion, rather than SVD, the difference in computation cost for performing the least-squares minimization is negligible. An example of the result that the transformation parameters part has on the image registration technique is seen in Figure 2.10. Figure 2.10a and Figure 2.10b are two aerial images of different orientation and rotation which contain partial overlapping areas. Through the use of the image registration algorithm, including the transformation parameters, the final image shown in Figure 2.10c is produced.

2.5.1 Iterative Weighted Least-Squares Algorithm

Using a similar method from robust statistics [53], the weighting matrix $W$ is used to distinguish between inliers and outliers. Applying the weights to the feature point pairs during each iteration will allow for a solution of the least squares problem in (2.14) that correspond to only the strongest matches. The weights are determined using the residual between the feature point pairs in the reference image ($B$) and the transformed feature point pairs using the updated least-squares solution ($AZ$). Thus small residuals will be heavily weighted, and large residuals, which indicate a non-matching pair, will have weights close to zero and thus have very little effect on the least square solution, or will be ignored completely. The algorithm for solving the least-squares problem is shown in Algorithm 1.
Figure 2.10: The following images show the overall effect of the image registration technique after the transformation parameters are obtained. With these transformation parameters, the two images can be combined using the corresponding points as the method to fit the two images together.
Algorithm 1: Iterative Weighted Least-Squares Minimization

1. Find an initial estimate of the transformation parameters $z$ by using the SVD with weights $w_i^{(1)} = 1, i = 1, 2, \ldots, K$

2. repeat

3. Compute the residuals for each feature pair point.

   \[ \Delta_i^{(n)} = \| f(P_i', z) - P_i \|^2, i = 1, 2, \ldots, K \]

4. Update the weights based on the values of the residuals using a robust estimator from (2.15).

5. Find a new solution for $z$ using (2.14) and the updated weights.

6. until $|\Psi^{(n)}(z) - \Psi^{(n-1)}(z)| < \epsilon$ or a maximum number of iterations has been reached.

The removal of outlier data is difficult to obtain on the first iteration, as the actual feature point pairs that present correspondence may be initially removed, hence the need for an iterative approach. The robust estimator predominately used in this method is an estimation which uses the median value of the residuals. Using the median, the robust estimation will ignore 50% of the residuals, which in some cases would filter out some inliers data that is above this value. This causes a slower convergence rate compared to using an estimation method based on the mean, however, due to the potential for a large number of outlier data, this method is well known to perform much stronger than the mean estimation [53]. In this application the robust estimator used is a popular form of $M$-Estimation proposed by Huber [54], which is referred to as Huber M-Estimation in this thesis. It is given by:

\[
w_i^{(n)} = \begin{cases} 
0 & \text{if } \Delta_i^{(n-1)} \geq \Lambda \\
1 - \left( \frac{\Delta_i^{(n-1)}}{Med(\Delta_i^{(n-1)})} \right)^2 & \text{otherwise}
\end{cases} \]  

(2.15)

\[
\Lambda = 1.354 \times 1.48 Med\left( \left| \sqrt{\Delta^{(n-1)}} - Med(\sqrt{\Delta^{(n-1)}}) \right| \right)
\]
Once the iterative weighting algorithm has been performed, a match can be determined by evaluating the magnitude of the error found between the transformed points in image $I$ and $\hat{I}$, from (2.13). The effect that the weighting function from (2.15) has is that it will exclude the outlier data from the error calculation, which provides an estimate of the fit between the two images using only the feature points that have correspondence.

### 2.6 Examples

Using the image registration algorithm discussed in this chapter, some examples are presented which show the successful application of image registration with various amounts of overlap, affine distortions and scale differences. Each example shows the two images that were used for the registration with their feature points overlaid on them, shown as yellow circle where size of the circle corresponds to the value of $s_{p_i}$. The feature points that were found using the correspondence of Zernike moments and output from the iterative weighted least squares minimization are shown as red crosses. The final registered image is shown as well as a plot showing the fit of the inlier and outlier feature points after the transformation parameters are applied to the target image and overlaid on the reference image. To evaluate the accuracy of the image registration algorithm, the error between the feature points of the transformed image and the reference image is calculated. This error is the distance, $D$, between the inlier feature point pairs, which is represented at the $\ell_2$ norm.

$$D = \| w_i \left( B_i - A_i z \right) \|$$

This distance error can be divided into the mean and standard deviation for all inlier feature point pairs [44].

$$D_M = \frac{1}{K} \sum_{i=1}^{K} D_i$$

$$D_{STD} = \sqrt{\frac{1}{K} \sum_{i=1}^{K} (D_i - D_M)^2}$$
where $K$ is the total number of feature point pairs after the iterative weighted least-squares algorithm.

The first demonstration features high resolution images acquired from Ikonos earth observation satellite [55]. Figure 2.11, which are aerial photographs taken of the University of California Santa Barbara campus, presents two images which have partial overlap with unity scale, that is to say there are no scale differences between the two images. The choice of scale values $s_{p_i}$ is arbitrary as the ratio that results in the best performance is 1. Figures 2.11a and 2.11b show the results from the feature point extraction and correspondence step of the image registration algorithm. The feature points extracted from each image and are shown as yellow circles on each image. The corresponding feature points between the two images that are found to match are shown as red crosses on Figures 2.11a and 2.11b. Estimating the transformation parameters of these corresponding feature point pairs results in the target image being transformed onto the reference image, shown in figure 2.11c. Figure 2.11d shows the correspondence of the feature point pairs after transformation, where one image has its feature points shown as squares while the other are crosses. The correct correspondences are shown as having the crosses coinciding with the squares. The outlier data which was eliminated by the iterative weighted least-squares optimization is also shown in Figure 2.11d in the form of black squares and crosses. It should be noted that there also appear to be correct corresponding points that are removed by the least-squares solution. This may be caused by the use of the Huber M-estimator in (2.15). For situations where a large number of corresponding feature point pairs are found, there may be some inlier points that will not be used. The error from the image registration algorithm ($D_M$ and $D_{STD}$) is 0.40 and 0.20 respectively, which shows an excellent registration between the two images.

Figure 2.12 consists of two images from the Landsat Thematic Mapper (TM)[56]. The extracted feature points and corresponding matching feature point pairs are shown on Figure 2.12a and 2.12b. Using the corresponding feature point pairs found by comparing their Zernike moments to all other pairs, the estimated transformation parameters can be obtained. Figure 2.12c shows the transformed target image superimposed on the reference
2.6 Examples

Figure 2.11: Registration of high resolution Ikonos images of University of California Santa Barbara [55].
2.6 Examples

Figure 2.12: Landsat Thematic Mapper (TM) Band 0 and Band 8 images. (a) and (b) The two images with the extracted feature points and corresponding matching feature points superimposed on each image. (c) The registered target image superimposed on the reference image. (d) The corresponding matching feature points for the target image (squares) superimposed on the reference image (crosses).
images. The transformed feature points are superimposed on the reference feature points in Figure 2.12d. It is observed that only the correct corresponding points remain after the iterative weighted least-squares minimization, shown as the colored squares (representing the target image) and crosses (representing the reference image) while the outlier data is shown as black squares and crosses. The error from the image registration algorithm ($D_M$ and $D_{STD}$) is 0.52 and 0.24 respectively, which shows an excellent registration between the two images.

An example using images showing differences in scale is provided in Figure 2.13. These images provide a demonstration where the color intensity is different while also having a significant scale change between the target and reference image. This image registration technique only uses the luminance of the image, reducing the effect of the differences in color. Performing the feature point extraction using a range of different scale parameters, $s_p$, results in the correspondence of matching feature point pairs. These points are superimposed on Figure 2.13a and 2.13b. After estimating the transformation parameters, the target image is transformed and superimposed on the reference images, shown in Figure 2.13c. The transformed feature points of the target image are superimposed on the reference feature points in Figure 2.13d. The error from the image registration algorithm ($D_M$ and $D_{STD}$) is 0.74 and 0.42 respectively, which shows a very good registration between the two images.

Figure 2.14 uses two images that are acquired from Google Earth. The target image, Figure 2.14b, has had scale and shearing distortion applied to it. Figure 2.14a and 2.14b show the extracted feature points and corresponding matching feature points pairs superimposed on the reference and target images. In this example, a larger number of scale ratios were required for filtering. The iterative weighted least-squares optimization was able to determine the transformation parameters required to register the target image with the reference image, shown in Figure 2.14c. Figure 2.14d shows the superimposed feature point pairs of the inlier and outlier data. This figure shows that there is a large number of

3Images are available at http://www.pbase.com/ckuhn55/wyoming/
Figure 2.13: Mountain Landscape images which has differences in color and scale
Figure 2.14: Aerial image of Paris, scale differences and affine distortions.
2.6 Examples

outlier feature point pairs that are rejected by the iterative weighted least-squares minimization. These outlier points can be due to self-similar data between the two images. Using the median value for the robust estimator allows the minimization algorithm to handle large amounts of outlier points. The error from the image registration algorithm ($D_M$ and $D_{STD}$) is 1.24 and 1.22 respectively, which shows a good registration between the two images.

The example in Figure 2.15 illustrates the case where full overlap between the target image and reference image exists. There also exists differences in scale and rotation between the two images, while the target image, Figure 2.15b, has gaussian noise applied to it. The extracted feature points and corresponding matching feature point pairs are superimposed on Figure 2.15a and 2.15b. The registered image in Figure 2.15c shows the target image transformed and superimposed on the reference image. Observing the location of the transformed feature points superimposed on the reference feature points after the iterative weighted least-squares minimization in Figure 2.15d shows that there were many outlier feature point pairs rejected. The error from the image registration algorithm ($D_M$ and $D_{STD}$) is 1.57 and 1.11 respectively, which shows a good registration between the two images.

These examples are compared using the methods from [11, 41, 45, 57] where applicable. For a comparison of techniques where scaling between the images are not present, such as [11, 57], the images from Figure 2.11 and 2.12 were used. The results from [11] were compared using the root mean squared error, RMSE, between the transformed feature points and the reference images feature points. The method from [57] uses a publically available registration tool called imREG. The results using the image registration algorithm from [45] are also provided to show how the changes made to the image registration algorithm in this chapter have on the performance of the image registration technique. The comparison of these techniques are shown in Table 2.1. Based on these results, it is seen that the image registration technique proposed in this chapter provide very good results when compared to other methods. Even comparing the difference between the original image registration technique from [45] and the improved algorithm in this chapter, the results
2.6 Examples

(a) First image with feature points  
(b) Second image with feature points  
(c) Registered image  
(d) Display of the fit between feature points using transformation parameters

**Figure 2.15:** Urban Landscape, scale differences, affine distortions and noise effects
are very good.

| Table 2.1: RMSE (in pixels) for images from in Figure 2.11 and 2.12 |
|----------------------|-------------------|-------------------|-------------------|-------------------|
| Ikonos               | N/A               | 1.27               | 0.45             | 0.25               |
| Landsat TM images    | 0.61              | 1.14               | 0.14             | 0.21               |

For images where scale changes are present, such as Figure 2.13, the SIFT technique from [41] is compared as well as the image registration technique which this chapter is based on[45]. For these methods, the registration errors ($D_M$,$D_{STD}$) are compared in Table 2.2. The results presented here show that, when compared to the SIFT method, or results from [45], the method in this chapter is very accurate. Further, due to the implementation of the algorithm using C, the algorithm is significantly faster than the technique from [45].

| Table 2.2: Registration Errors (in pixels) for images from in Figure 2.13 |
|----------------------|-------------------|-------------------|-------------------|-------------------|
| Image                | Method from [41]  | Method from [45]  | Method from this Chapter |
| Mountain             | 0.40              | 0.19              | 0.86              | 0.46              | 0.74 | 0.42 |

2.7 Summary

This chapter presented an image registration algorithm which made use of scale interactions of Mexican-hat wavelets for feature point extraction, magnitudes of Zernike moments for correspondence and an iterative weighted least squares minimization algorithm to determine transformation parameters. The discussion of the effect that scale between images has on image registration and the importance of the ratio between scale parameters $s_{pi}$ of the feature extraction step has been discussed.

Examples were provided to illustrate the application of this image registration algorithm for various image such as satellite images and included affine distortions, scale difference
and noise distortions.
Chapter 3

Face Recognition

3.1 Abstract

This chapter proposes a technique of the image registration technique for the face recognition problem. The face recognition problem in this specific application is the matching between two subject images where a face is present. For the face recognition problem, in order to have success, the application is required to determine whether a genuine match is found when the subject images are the same face and when the subject images are of different faces.

The chapter is organized as follows; the motivation and purpose for using image registration in the application of face recognition is discussed in section 3.3. Section 3.4 outlines the proposed algorithm for face recognition and the adaptations of the image registration algorithm to face recognition is discussed in detail. An example of the proposed method in this chapter is performed with images varying in pose, illumination and background is shown in section 3.5. The performance of the method to the face recognition problem is discussed in the next chapter and will include comparisons with existing techniques.

3.2 Introduction

The image registration technique discussed in the previous chapter which is based on the scale interaction of Mexican-hat wavelets [16, 17] will be used for face recognition. This
method of image registration has shown to have a good performance with geometric affine transformations and scale differences between two images [17]. Feature points in the image of a face which are related to the change in curvature of the intensity of the image are obtained using filtering with Mexican-hat wavelets are then used to find the correspondence between images. The correspondence between feature points is based on a descriptor vector consisting of the magnitudes of the Zernike moments around the feature points. The parameters of an affine transformation between the sets of features points is obtained using an iterative least squares algorithm which helps to eliminate outliers and the recognition is based on the residual error of the inliers after transformation.

Face recognition provides an interesting application for image registration. Specifically, given the uniqueness of the human face to other objects, the image registration technique can be used to accurately find unique feature points of a given individual and compare them to other images. This application is used to specifically identify individuals in images, as opposed to general object recognition of faces. The choice of using the image registration technique for a proposed application of face recognition comes from the advantages of allowing images with partial overlap, differences in scale and affine distortions.

3.3 Application of Image Registration for Face Recognition

The face recognition problems poses a number of challenges, primarily with respect to changes in pose and illumination conditions. While the shape of the face is consistent and always has the same features, the rotation of the human head will not always create the same view in images. Furthermore, illumination of the surroundings to the human face may darken or enhance features, even the camera that is capturing the image may encounter over and underexposure, which can make the feature points difficult to detect. This means that the image registration algorithm has to be robust to changes in the position of the face
and also to changes in the luminance of an image.

Using the algorithm discussed in the previous chapter, an overall face recognition algorithm can be created which uses the image registration technique as the main component. These challenges have been addressed by creating a training set of images that consist of different poses and illuminations with the intent that this training set will be representative of any effects that are present in the environment. Taking into account that in most situations, the face of a subject is naturally normalized (that is the location of the face is generally upright and forward facing), a measureable criteria can be observed to ensure that the correct registrations are found. This decision criteria is determined from a post-processing step that observes the affine distortions found from the estimated transformation parameters. These additions to the image registration technique are explained in detail in section 3.4.
3.4 Mexican Hat Based Algorithm for Face Recognition

The algorithm for performing the face recognition is summarized in Figure 3.1. The feature extraction and registration part is based on the image registration technique discussed in the previous chapter and in [16, 17, 58]. It is performed by filtering an image with Mexican-hat wavelets which provides a relation to the change in curvature of the intensity of the image and can be used to identify features in the facial image. These features are then used to find the correspondence between images. This correspondence between feature points is based on a descriptor vector consisting of the magnitude of the Zernike moments around the feature points. The parameters of an affine transformation between the sets of features points are obtained using an iterative weighted least-squares algorithm which helps to eliminate outliers. The recognition is based on the evaluation of a post-processing step which uses the residual error of the inliers and the information from the affine transformation parameters. A training step which uses multiple images in the test set for performing comparisons to a subject image is used to help deal with possible differences in illumination and pose while also reducing the execution time of the algorithm. A post-processing step has been introduced to improve the decision making whether a genuine match has been found and to help remove false acceptances introduced by large outlier data. The details of the training and post-processing steps will be described in the rest of this section.

3.4.1 Training Step

The training step is meant to compensate for changes in pose and illumination differences by incorporating multiple images in the test set. The approach proposed in this chapter using multiple training images in connection with the registration method in [16, 17, 58] is different than the one proposed in [59] and [60] which is based on preprocessing, but has yielded better results. Furthermore, the execution time of the algorithm by using this training step is reduced due to performing the filtering and Zernike moment calculation of the training images once and storing the result in a data file.
Figure 3.2: Typical procedure for creating training data. Training images (or a single image) are selected that cover a range of different pose and illumination conditions and are processed by the feature extraction steps of the image registration algorithm. The resulting features and Zernike moments for each image are stored in data files.

The procedure for creating the training set of images consists of the following steps.

1. Choose training images.
2. Crop images to only contain the face region.
3. Perform feature extraction and Zernike moment calculation on each image.
4. Store the location of the feature points and the magnitude of the Zernike moments for each feature point into a data file.

In the first step the choice of training images to be used are subject to the type of conditions that are needed to be compensated against. For example, if the database of images only deals with pose changes, then the training set could include images of differing pose. If there are changes in both pose and illumination, then images from both conditions can be included. The number of training images to use is decided by the degree of change within the pose and/or the illumination condition. Images that contain a wide range of head movement or lighting changes would use more training images to help compensate. While this method puts emphasis on using multiple images, for simple comparisons with little change in pose and illumination, one image as a training image can still be used.
3.4 Mexican Hat Based Algorithm for Face Recognition

with good results. Once the training images are chosen, they are cropped closely to the face region to ensure that correspondences found are from facial information only and not due to background detail or similarities in clothing. The filtering, feature point extraction and calculation of Zernike moments are performed and stored in data files. Storing this information is done so that the algorithm is not repeating the same computations multiple times leading to faster execution time.

3.4.2 Post-Processing

To determine whether the transformation parameters that were estimated in registration algorithm provided a valid match, a post processing procedure was used. The post-processing has been modified to minimize these effects in skew, shear and rotation by observing 3 key parameters used in the decision criteria which are obtained from the results of estimating the affine transformation parameters.

1. Error parameter, $\alpha_{\text{error}}$.
2. Scaling parameter, $\alpha_{\text{scale}}$.
3. Rotation parameter, $\alpha_{\text{rot}}$.

The error parameter, $\alpha_{\text{error}}$, consists of the ratio between the residual error to the sum of all the weights from the weighted least squares minimization from the inlier data and is shown in (3.1). This parameter provides information as to the accuracy of the fit between the two images. This ratio indicates the size of the error relative to the number of inlier points determined from the iterative weighted least squares minimization. This provides a scaleable value which is based on the number of matching points found by the feature extraction and registration step of the algorithm. A low value of $\alpha_{\text{error}}$ shows an increased confidence in the matching result. A small number of matching points indicates a weak confidence on the matching result. For correct registrations, the RMSE error between the feature points of the two images will be small while for incorrect registrations will have
large RMSE values.

\[ \alpha_{\text{error}} = \sqrt{\sum w_i (\hat{x}_i - x_i)^2 / \sum w_i} \]  

(3.1)

The scaling criteria, \( \alpha_{\text{scale}} \), is determined by decomposing the estimated affine parameters using singular value decomposition and taking the ratio between the singular values, shown in (3.2).

\[
\begin{bmatrix}
\hat{x}_1 \\
\hat{x}_2
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} +
\begin{bmatrix}
t_{x1} \\
t_{x2}
\end{bmatrix}
\]

\( \hat{x}_i = Zx_i + T \)

\( Z = U\Sigma V^T \)

where the singular values are the elements of the diagonal matrix \( \Sigma \)

\[
\Sigma =
\begin{bmatrix}
\sigma_1 & 0 \\
0 & \sigma_2
\end{bmatrix}
\]

\[ \alpha_{\text{scale}} = \frac{\sigma_1}{\sigma_2} \]  

(3.2)

This provides information whether any non-uniform scaling is being performed between the two images (\( \alpha_{\text{scale}} \gg 1 \)). This non-uniform scaling is an estimate of the skew and shearing effects. For images that contain the same subject this ratio will be very close to unity, with slight increases or decreases to account for differences in pose. In images that have different subjects, this ratio is found to identify cases where the transformation parameters provide a small error but are geometrically incorrect. These cases can be rejected based on it which results in the correct rejection of these matches.

The rotation criteria, \( \alpha_{\text{rot}} \), is the estimated rotation of the affine parameters, this is found using the \( U \) and \( V^T \) matrices from the singular value decomposition in (3.2) which indicate an estimate of the pre/post-scaling rotation operations which is shown in (3.3).

\[
UV^T =
\begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\]

\[ \alpha_{\text{rot}} = \theta \]  

(3.3)
For the face recognition application, the post-processing should only allow the affine parameters to have slight changes in rotation and shearing distortions, as all face images have similar orientation. Observing that in cases where the subjects are the same, a relationship exists between the error parameter and scale parameter. For images where the two subject images are of the same face, the error parameter and scale parameter are both small. For images where there is a difference between the subject faces, the error and scale parameters are large. Figure 3.3 shows an example of such a distribution. From this relationship, a threshold value $\alpha_{\text{thresh}}$ can be created that represents a linear boundary between the area where a match occurs and a non-match occurs. This is given in (3.4). The values for the slope and y-intercept of (3.4) were found by observing the effect that different values of $\alpha_{\text{scale}}$ and $\alpha_{\text{error}}$ had on sample images. Manually adjusting these parameters and viewing the resulting fit between the two images provided a basis for adjusting the size of the boundary region. A triangular boundary was created to allow for larger shear and skew effects to be allowed for very low error parameters. Thus when the fit between the two images is very good, more shear and skew is tolerated. The same is true for very small non-uniform scaling. Where there is little shear and skew effect, the boundary allows for a larger error parameter.

$$\alpha_{\text{thresh}} = \frac{-2.5}{0.4} (\alpha_{\text{scale}} - 1) + 2.5$$

(3.4)

The decision whether two face images match is determined based on the analysis of these parameters and is summarized in (3.5). The conditions for a valid match must be found such that the angle parameter has less then $7.5^\circ$ rotation and the error parameter is less then the calculated error threshold from (3.4). If either of these conditions are false then the face recognition algorithm considers the images are not a match.

$$\text{RESULT} = \begin{cases} 
\text{MATCH} & \alpha_{\text{error}} \leq \alpha_{\text{thresh}} \quad \text{and} \quad \alpha_{\text{rot}} < 7.5^\circ \\
\text{NO MATCH} & \text{otherwise}
\end{cases}$$

(3.5)
Figure 3.3: Plot showing the distribution between the error and scale parameter. For subjects that are the same face, the distribution is close around the origin. For different faces between the subjects, the distribution is largely varied. Comparisons of images with the same subject are represented with ×. Comparisons of images of different subjects are represented with o.
3.5 Demonstration of the Proposed Method

The technique proposed in this chapter can be demonstrated with a general case of comparing a cropped training image against two images, one of the same subject and one of a different subject. A demonstration is conducted for three distinct scenarios; differences in pose, differences in illumination and differences in background and scale.

3.5.1 Demonstration Against Cases of Varying Pose

The demonstration of the proposed face recognition algorithm for changes in pose is shown in Figure 3.4. A training step is done where the test image is cropped about the face area to ensure that correspondence is not falsely linked to the background (Figure 3.4a). Two sample images are tested against, one of the same subject (Figure 3.4b) and one of a different subject (Figure 3.4c). The feature extraction and registration step is performed which determines the matching feature points corresponding to both the training image and the sample image (Figure 3.4d, Figure 3.4e). Finally the post-processing decision criteria are calculated which determine whether a match has been found.

For the case of the same subject in training and sample images, the obtained parameters used for the decision criteria are:

\[ \alpha_{\text{error}} = 0.12966 \quad \alpha_{\text{scale}} = 1.006139 \quad \alpha_{\text{rot}} = 1.6743^\circ \]

Based on these parameters, error threshold, \( \alpha_{\text{thresh}} \) was calculated to be 2.4616, and thus, from the rules for a match in (3.5), a match was found. The result is shown in Figure 3.4f.

For the case where the training image and sample image contains different subjects, the obtained parameters used for the decision criteria are:

\[ \alpha_{\text{error}} = 4.0382 \quad \alpha_{\text{scale}} = 1.18852 \quad \alpha_{\text{rot}} = 18.1153^\circ \]

Based on these parameters, it is found that the error threshold, \( \alpha_{\text{thresh}} \) is found to be 1.3218. Based on the rules for a match in (3.5), a non-match was found. This was due to both the error criteria and rotation criteria being rejected. The result is shown in Figure 3.4g.
Figure 3.4: Demonstration of the proposed algorithm versus pose changes. (a) The image used for the training data is cropped about the face. (b) Image of the same subject with different pose. (c) Image of a different subject with different pose. (d)(e) Matching points found from the feature extraction and registration step overlaid. (f) Training image overlaid on the subject image showing a matching registration. (g) Training image overlaid on the subject image showing a non-matching registration.
3.5 Demonstration of the Proposed Method

3.5.2 Demonstration Against Cases of Varying Illumination

An example demonstrating changes in illumination is shown in Figure 3.5. Similar to the example versus pose, the cropped image of the training face are shown in Figure 3.5a. Two images, one of the same subject and one that is different, are chosen that have a change in illumination condition from the training image (Figure 3.5b, 3.5c), these images show much more shadow then that of the cropped image. The resulting matching points from the feature extraction and registration step on the sample images are shown on each sample image (Figure 3.5d, 3.5e).

For the same subject in both the training and sample images, the post-processing step obtained the following parameters used for the decision criteria:

$$\alpha_{\text{error}} = 0.2658 \quad \alpha_{\text{scale}} = 1.14193 \quad \alpha_{\text{rot}} = 2.8601^\circ$$

From these parameters the error threshold, $\alpha_{\text{thresh}}$ was calculated as 1.6129. This indicates that a match was found between the training image and subject image. Figure 3.5f shows the resulting images overlaid.

Observing the case where the training image and sample image contains different subjects, the obtained parameters used for the decision criteria are:

$$\alpha_{\text{error}} = 6.3773 \quad \alpha_{\text{scale}} = 2.4378 \quad \alpha_{\text{rot}} = 58.353^\circ$$

Based on these parameters, it is found that the error threshold, $\alpha_{\text{thresh}}$ is found to be $-6.4863$, and thus, based on the rules for a match in (3.5), a non-match was found. Figure 3.5g shows the resulting effect that the transformation parameters from the feature extraction and registration step found, which is clearly not a correct registration of two faces.

3.5.3 Demonstration Against Cases of Varying Background and Scale

To illustrate the effect of the face recognition algorithm with respect to changing background locations, scale and posture/facial expression is shown in Figure 3.6. Prior to the
Figure 3.5: Demonstration of the proposed algorithm versus illumination changes. (a) The image used for the training data is cropped about the face. (b) Image of the same subject with different pose. (c) Image of a different subject with different pose. (d)(e) Matching points found from the feature extraction and registration step overlaid. (f) Training image overlaid on the subject image showing a matching registration. (g) Training image overlaid on the subject image showing a non-matching registration.
3.5 Demonstration of the Proposed Method

Figure 3.6: Demonstration of the proposed algorithm versus changing background and scale changes. (a) The image used for the training data is cropped about the face. (b) Image of the same subject with different pose. (c) Image of a different subject with different pose.

start of the algorithm, a training image is cropped to only include the face (Figure 3.6a) and has it’s feature point locations and magnitudes of corresponding Zernike moments stored. Two images are chosen, one of the same subject and one of a different subject (Figure 3.6b and 3.6c), which have changes in backgrounds, positioning, posture, scale and facial expression. The matching feature points found from the feature extraction and registration step of the algorithm are then overlaid to show where the correspondences are found (Figure 3.6d and 3.6e).

Using the example where both the training image and subject image are of the same subject, the obtained parameters used for the decision criteria are:

\[
\alpha_{\text{error}} = 0.047106 \quad \alpha_{\text{scale}} = 1.02938 \quad \alpha_{\text{rot}} = 0.14151^\circ
\]

The resulting error threshold, \(\alpha_{\text{thresh}}\), is calculated to be 2.3164 which corresponds to a match. The resulting registration between the training image and subject image shows the result of the successful matching registration in Figure 3.6f.

In the case where the training image and sample image contain different subjects, the obtained parameters used for the decision criteria are:

\[
\alpha_{\text{error}} = 2.0746 \quad \alpha_{\text{scale}} = 20.1983 \quad \alpha_{\text{rot}} = 84.7501^\circ
\]
3.5 Demonstration of the Proposed Method

Figure 3.6: Demonstration of the proposed algorithm versus changing background and scale changes. (d)(e) Matching points found from the feature extraction and registration step overlaid. (f) Training image overlaid on the subject image showing a matching registration. (g) Training image overlaid on the subject image showing a non-matching registration.
Based on these parameters, it is found that the error threshold, $\alpha_{\text{thresh}}$, is found to be $-117.4896$, and thus, based on the rules for a match in (3.5), a non-match was found. Specifically observing the value of $\alpha_{\text{scale}}$ it is seen that there is a large amount of skewing in the image which is not consistent with the expected orientation of a face object. Figure 3.5g shows the resulting effect that the transformation parameters from the feature point extraction and registration step found, which is clearly not a correct registration of two faces as the sample image has be severely distorted in order to geometrically fit the feature points.

### 3.6 Summary

This chapter discussed the application of image registration for use in the face recognition problem. An application for face recognition was proposed using the image registration technique from [16, 17, 58] which uses the scale interactions of Mexican-hat wavelets for the feature extraction. The Zernike moments from these features points are used in an iterative weighted least squares minimization for determining the correspondence. Using this image registration technique, modifications can be made to this algorithm to allow for the compensation of differences in pose and illumination conditions and also to help determine the success of the correspondence between the two images. These modifications are in the form of using multiple training images and in a post-processing step that analyzes the result of the iterative weighted least-squares minimization to obtain a set of parameters that are used for the decision criteria to determine if a correct match is found.

Examples illustrating the proposed face recognition technique was demonstrated using images of varying pose, illumination and background. These demonstrations showed the successful matching and rejection of images containing faces. In cases where the training image and subject images were from the same subject, it was shown that the proposed algorithm was able to identify common features and determine that a genuine match as found. In the case were the training image and subject image were of different subjects, the
proposed algorithm was able to use the results from the post-processing step to determine that the resulting transformation parameters did not provide a consistent match between the two images.
Chapter 4

Experimental Design and Results

4.1 Abstract

In this chapter the performance of the proposed method for face recognition is evaluated using databases of images from various experiments. The experimental design, data used, results and comparison to existing techniques are discussed. The chapter is organized as follows: Section 4.3 discusses the use of the Yale Face Database B [29] for experiments in changing pose and illumination. Section 4.4 will discuss the experiments using the Caltech Face Database [42, 43] for an experiment on the effect of changing background, scale, pose, illumination and expression. Conclusions from these experiments are summarized in section 4.5.

4.2 Introduction

This chapter provides a description of the experiments for the proposed face recognition algorithm using image registration. The results of these experiments are compared to existing methods that made use of the same data set for an accurate measure in performance of the proposed method. The methods that are used for comparison include methods that are appearance-based (PCA, ICA, etc) and also a method that creates 3D models for comparison[25, 27, 29].

For the two groups of experiments that are used in this section, each is an example of an
application for face recognition. The first set of experiments uses a database of images generated in a very controlled environment where the pose and illumination of all the images is varied uniformly throughout the database. Such an example would be similar to using government identification images (drivers license, passport photos, criminal identification photos). The purpose of using this data for the first experiment is to test the performance of the proposed method to specific changes in pose and illumination conditions. The second experiment in this section makes use of a database where the images are taken in a largely uncontrolled environment. This database resemble images that are taken in a social setting. The background of the images are dynamic, and changing between subjects. The position of the camera and the subject location in each image is different for each image which creates subtle changes in scale and orientation. Furthermore, some images contain changes in facial expression. An application using images similar to this database could be used in social networking applications, multimedia search, and content based image retrieval applications.

The performance of the algorithm can be evaluated using the following criteria: the total error rate, which is a total of both the total false rejection rate and total false acceptance rate is calculated using (4.1). To accurately show the effect that false rejection and false acceptance are having on the experiment results a per subject and per non-subject error rate is also calculated using (4.2a) and (4.2b). A non-subject is defined as all images that are not of the same person as the target image in the training data.

\[
\begin{align*}
FR_{Total} &= \frac{\sum \text{False Rejections}}{\text{Total Images}} \\
FA_{Total} &= \frac{\sum \text{False Acceptances}}{\text{Total Images}} \\
\text{Total Error} &= \frac{\sum \text{False Rejections} + \sum \text{False Acceptances}}{\text{Total Images}} \\
\end{align*}
\] (4.1)

\[
\begin{align*}
FR_{Subject} &= \frac{\sum \text{False Rejections}}{\text{Total Images per Subject}} \\
FA_{Subject} &= \frac{\sum \text{False Acceptances}}{\text{Total Images per Non-Subject}} \\
\end{align*}
\] (4.2a) (4.2b)
4.3 Experiments using the Yale Face Database B

The experiments to evaluate the performance of the proposed face recognition algorithm to changes in pose and illumination use the images provided by the Yale Face Database B [29]. This database was specifically created to provide a controlled set of high-resolution images of various subjects under differing conditions of pose and illumination. In [29] a specially designed geodesic dome was created to capture images of subjects from different camera angles and under different light source locations. The database is organized into specific
4.3 Experiments using the Yale Face Database B

subsets based on pose and illumination, which are shown in Figure 4.1. Each column of Figure 4.1 represents a subset of images with a particular illumination as follows:

- **Subset 1:** Illumination angles $\theta < 12^\circ$ from optical axis
- **Subset 2:** Illumination angles $20^\circ < \theta < 25^\circ$ from optical axis
- **Subset 3:** Illumination angles $35^\circ < \theta < 50^\circ$ from optical axis
- **Subset 4:** Illumination angles $60^\circ < \theta < 77^\circ$ from optical axis

Each of these subsets contains images corresponding to different poses as follows:

- **Frontal:** Camera position in front of subject (1 image)
- **12°:** Camera position at 12° difference from frontal pose (5 images)
- **24°:** Camera position at 24° difference from frontal pose (3 images)

In total the Yale Face Database B consists of 10 unique subjects, each of which have 9 poses (1 frontal pose, 5 poses offset by 12° and 3 poses offset by 24°) and 44 illuminations (6 in subset 1, 12 in subset 2, 12 in subset 3 and 14 in subset 4) for a total database size of 3960 images.

Three specific experiments are performed using this database and followed the experimental design from [29]. Experiment 1 performs a test on the effect that illumination has on the face recognition accuracy. This is performed using the same frontal pose for all images in the illumination subset 2,3 and 4. The size of the database for this experiment is a fraction of the total database size, it consists of 380 images (38 illumination images/subject, 10 subjects). Experiment 2 is a test on the effect that pose has on the face recognition accuracy, which is found for all images in three the pose subsets within the same illumination subset. This experiment uses all the images within illumination subset 2, which consists of 540 images (6 images/pose, 9 poses/subject, 10 subjects). Experiment 3 focuses on testing the face recognition accuracy using all the images in pose and illumination subsets. This experiment uses the full size of the database.

The experiments were conducted using the proposed face recognition algorithm proposed in chapter 3. The training and postprocessing steps were used as described in sec-
Table 4.1: Average comparison time in seconds per image

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>Average Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Image</td>
<td>1.2706</td>
</tr>
<tr>
<td>3 Images</td>
<td>1.3192</td>
</tr>
<tr>
<td>4 Images</td>
<td>1.3424</td>
</tr>
<tr>
<td>6 Images</td>
<td>1.3994</td>
</tr>
</tbody>
</table>

4.3 Experiments using the Yale Face Database B

To test the affect that the addition of using multiple training images has, the following four specific scenarios were evaluated:

- **1 image**: 1 Frontal pose from illumination subset 1.
- **3 images**: 1 Frontal pose, 2 poses offset by $12^\circ$, all from illumination subset 1.
- **4 images**: 2 Frontal pose one each from illumination subset 1 and 3; 2 poses offset by $12^\circ$ from illumination subset 1.
- **6 images**: 2 Frontal pose one each from illumination subset 1 and 3; 4 poses offset by $12^\circ$, two from illumination subset 1 and two from subset 3.

To determine the necessary scale parameters and error parameters for the feature extraction and post-processing steps respectively, a test set of images was evaluated. The test set was chosen to consist entirely from illumination subset 1, shown as the first column in Figure 4.1. Because of the small amount of images of subset 1, 6 images per pose for each subject for a total of 540 images, it provided a small sample of images to determine these parameters. The distribution of the results for all the images in this learning set were shown in Figure 3.3 from the previous chapter. In order to validate the face recognition algorithm the remaining images from subsets 2, 3, and 4 were used as a validation set for the experiments using this database.

The execution of the face recognition algorithm was implemented using MATLAB. Matlab executable (MEX) functions were created using C to handle the feature extraction and Zernike moment correspondence routines. Table 4.1 shows the average execution time...
to perform a comparison between the training and the sample images.

4.3.1 Experiment 1: Effect of Varying Illumination with Constant Pose

The first experiment performed is a test between varying illumination conditions while maintaining a fixed frontal pose. To perform this experiment, images from the training scenarios above were compared to all frontal images in subsets 2, 3 and 4 for all 10 subjects. Each face in a training image was found manually and extracted to ensure that only the face area of the image was being compared to prevent false acceptance due to background similarities. The sample size for this experiment is 380 images, 38 images per subject of frontal pose within illumination subset 2, 3 and 4 between 10 subjects.

The error rates found in Table 4.2 shows the accuracy of the face recognition algorithm with respect to false rejection, false acceptance and total error. Table 4.3 shows the breakdown of the per subject error rate from (4.2). From these results it is observed that the proposed face recognition algorithm performed reasonably well in terms of maintaining a low false rejection rate for changes in illumination, however as the illumination angle increased (shown in the difference between subsets) the false rejection rate increased significantly. The addition of more images to the training step helped to reduce the effect that illumination had on the subject image. Specifically when observing the difference between one training image and four training images where the error rate is much improved. The significant reduction in this error rate is seen in the improvement of the false rejection rate, with only a small increase seen in the false acceptance rate. While the overall error rate is reduced, it is noted that for 6 training images, the false rejection rate increased significantly over the 4 image case. This is partially because in the 6 image case, the decision of if a match is found is based on a 2 out of 6 voting. Meaning that two of the trained images in the training set must record a match in order for a proper match to be found. Whereas the 4 image training set uses a 1 out of 4 voting. While this increases the false rejection rate, it
4.3 Experiments using the Yale Face Database B

significantly improves the false acceptance rate.

**Table 4.2:** Total Error summary for Experiment 1

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>Error Rates (%)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FR&lt;sub&gt;total&lt;/sub&gt;</td>
<td>FA&lt;sub&gt;total&lt;/sub&gt;</td>
<td>Total Error</td>
</tr>
<tr>
<td>1 Image</td>
<td>6.05</td>
<td>0.26</td>
<td>6.32</td>
</tr>
<tr>
<td>4 Images</td>
<td>2.89</td>
<td>1.05</td>
<td>3.94</td>
</tr>
<tr>
<td>6 Images</td>
<td>5.00</td>
<td>0.00</td>
<td>5.00</td>
</tr>
</tbody>
</table>

**Table 4.3:** Breakdown of the False Acceptance and False Rejection rate for Experiment 1

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>Error Rates (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FR&lt;sub&gt;subject&lt;/sub&gt;</td>
<td>FA&lt;sub&gt;subject&lt;/sub&gt;</td>
</tr>
<tr>
<td>1 Image</td>
<td>60.53</td>
<td>0.29</td>
</tr>
<tr>
<td>4 Images</td>
<td>28.95</td>
<td>1.17</td>
</tr>
<tr>
<td>6 Images</td>
<td>50.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Due to the much smaller size of subject images compared to non-subject images in the database, these small increases in the false rejection errors show up as large errors in the false rejection rate, FR<sub>subject</sub>.

### 4.3.2 Experiment 2: Varying Pose and Constant Illumination Condition

The second experiment performed was tested using varying pose while the illumination condition is kept constant. This experiment was performed by comparing the training images to all the images within illumination subset 2. The choice to use illumination subset 2 was because the change in illumination between this subset and the training images is minimal. Within illumination subset 2, there are a total of 12 different images available for each of the 9 poses. In total this experiment uses 1080 images from the Yale Face Database B.
The results from this experiment are shown in Table 4.4. These results are broken into false rejection and false acceptance rate to show the distribution of the error rate. Table 4.5 further breaks the false rejection and false acceptance rate to show the distribution of the error within each subset of subject and non-subject images from (4.2a) and (4.2b).

### Table 4.4: Total Error summary for Experiment 2

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>FR&lt;sub&gt;total&lt;/sub&gt;</th>
<th>FA&lt;sub&gt;total&lt;/sub&gt;</th>
<th>Total Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Image</td>
<td>4.63</td>
<td>0.65</td>
<td>5.28</td>
</tr>
<tr>
<td>3 Images</td>
<td>1.30</td>
<td>1.11</td>
<td>2.41</td>
</tr>
</tbody>
</table>

### Table 4.5: Breakdown of the False Acceptance and False Rejection rate for Experiment 2

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>FR&lt;sub&gt;subject&lt;/sub&gt;</th>
<th>FA&lt;sub&gt;subject&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Image</td>
<td>46.30</td>
<td>0.72</td>
</tr>
<tr>
<td>3 Images</td>
<td>12.96</td>
<td>1.23</td>
</tr>
</tbody>
</table>

From the results in Table 4.4 provide a very good result for conditions of changing pose. Adding training images is shown to improve the performance of the experiment on changing pose. In the case where 3 training images are used, a significant improvement in the error rate is shown. From Table 4.5, the false rejection rate shows a large improvement while the false acceptance rate is only slightly increased.

Table 4.6 shows the breakdown of the results based on the different pose in each subset. It is seen that the proposed method works well for images of slight changes in pose. This can be attributed to the performance of using the Zernike moments of the feature points as well as the iterative weighted algorithm for removing the outliers, which allow for a close match between the detected features.
4.3 Experiments using the Yale Face Database B

Table 4.6: Experiment Results with Constant Illumination with Varying Pose (Experiment 2)

<table>
<thead>
<tr>
<th>Method</th>
<th>Error Rates (%) vs. Pose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frontal</td>
</tr>
<tr>
<td>1 Image</td>
<td>3.33</td>
</tr>
<tr>
<td>3 Image</td>
<td>3.33</td>
</tr>
<tr>
<td>6 Image</td>
<td>0.00</td>
</tr>
</tbody>
</table>

4.3.3 Experiment 3: Effect of Varying Illumination and Varying Pose

The third experiment performed a test against variations in both pose and illumination. The database of images consisting of all the poses within illumination subsets 2, 3 and 4 are used, totalling in 3420 images. This experiment was executed by taking the training set images and comparing them to all poses contained in illumination subsets 2, 3 and 4 for all 10 subjects. The results are compared to those of the same experiment from [29] later in this chapter.

The results from this experiment are shown in Table 4.7. These results are categorized into false rejection and false acceptance rate to show the distribution of the total error rate. Table 4.8 further breaks down the false rejection and false acceptance rate to show the distribution of the error between subject and non-subject images from (4.2a) and (4.2b).

Table 4.7: Total Error summary for Experiment 3

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>Error Rates (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FR&lt;sub&gt;total&lt;/sub&gt;</td>
</tr>
<tr>
<td>1 Image</td>
<td>8.04</td>
</tr>
<tr>
<td>3 Images</td>
<td>6.32</td>
</tr>
<tr>
<td>4 Images</td>
<td>5.23</td>
</tr>
<tr>
<td>6 Images</td>
<td>6.81</td>
</tr>
</tbody>
</table>

From the results in Table 4.7, the face recognition algorithm provides a very good result for conditions of changing pose and illumination. Adding more training images that
4.3 Experiments using the Yale Face Database B

Table 4.8: Breakdown of the False Acceptance and False Rejection rate for Experiment 3

<table>
<thead>
<tr>
<th>Number of Training Images</th>
<th>Error Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FR_{subject}</td>
</tr>
<tr>
<td>1 Image</td>
<td>80.41</td>
</tr>
<tr>
<td>3 Images</td>
<td>63.16</td>
</tr>
<tr>
<td>4 Images</td>
<td>52.34</td>
</tr>
<tr>
<td>6 Images</td>
<td>68.13</td>
</tr>
</tbody>
</table>

address pose and illumination conditions, as in the case of 4 images and 6 images scenario is shown to help reduce the false rejection rate with only a slight cost of an increased false acceptance rate, while the total error rate is reduced. Similar to experiment 1, the results for the 6 training image scenario are slightly higher, however, due to the 2 out of 6 voting decision that is used in this case, the false rejection rate is increased. Whereas the false acceptance rate is significantly reduced compared to the other methods.

4.3.4 Comparison of Results With Other Algorithms

Results from other methods reported in [29] were used for comparison with the proposed face recognition algorithm. The first method used in [29] is a simple nearest neighbor matching classification [27]. Simply, this method will assign an image from the test set to the lowest distance based match from the learning set images. Such an approach was named Correlation when referring to images that are normalized for zero mean and unit variance. A principal component analysis (PCA) approach was also used, which is also commonly referred to as Eigenfaces [25]. A modified version of Eigenfaces which removes the first three most significant principal components and having been demonstrated to have improved results over the original method [25] was used in [29]. The illumination cone methods [29] were also used for a comparison. This technique is an example of one that is designed for robustness to illumination and pose.

The results for the first experiment where the effect of changing illumination conditions
was performed were compared with the results of other methods in [29] and shown in Table 4.9. To ensure consistency with these results, only the total error rate for the proposed algorithm is shown. This displayed error rate was interpreted from the results of [29] as it was not explicitly stated. These results show that the proposed face recognition algorithm performs well compared to the current PCA techniques, although not as good as the illumination cone method presented in [29]. The proposed algorithm was noted to run at a very fast execution time for the comparison of two images, which was measured and shown in Table 4.1. No information was available as to the speed for the other algorithms, in particular the illumination cone which requires this particular algorithm to make use of creating a 3D model of the subject, which is expected to be computationally expensive.

Table 4.9: Comparison of Results with Frontal Pose and Varying Illumination Conditions (Experiment 1)

<table>
<thead>
<tr>
<th>Method</th>
<th>Error Rate (%) vs. Illumination</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subset 2</td>
</tr>
<tr>
<td>Correlation[27]</td>
<td>N/A</td>
</tr>
<tr>
<td>Eigenfaces[23]</td>
<td>N/A</td>
</tr>
<tr>
<td>Eigenfaces[25] w/o 1st 3</td>
<td>N/A</td>
</tr>
<tr>
<td>Linear Subspace[28]</td>
<td>N/A</td>
</tr>
<tr>
<td>Cones-attached[29]</td>
<td>N/A</td>
</tr>
<tr>
<td>Cones-cast[29] (Subspace Approx.)</td>
<td>N/A</td>
</tr>
<tr>
<td>Cones-cast[29]</td>
<td>N/A</td>
</tr>
<tr>
<td>1 Image</td>
<td>0.83</td>
</tr>
<tr>
<td>4 Images</td>
<td>0.00</td>
</tr>
<tr>
<td>6 Images</td>
<td>0.83</td>
</tr>
</tbody>
</table>

The results for the second experiment, where the pose was varied with constant illumination, had no data for comparison from other reference material, and thus no comparison can be shown.
The results of the comparison for the third experiment are listed in Table 4.10. Note that these results are grouped based on pose and contain images from all illumination subsets in the database. From this comparison, it is shown that the proposed face recognition algorithm has a consistent performance when adjusting the pose and the illumination angles, and similar to the results from the first experiment, the proposed algorithm performs very well compared to existing PCA techniques, but not as good as the illumination cone methods. As expected, the proposed method performed best for the frontal position.

Table 4.10: Comparison of Results for Varying Pose and Varying Illumination using subsets 2, 3 and 4. (Experiment 3)

<table>
<thead>
<tr>
<th>Method</th>
<th>Error Rates (%) vs. Pose</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frontal</td>
</tr>
<tr>
<td>Correlation[27]</td>
<td>23.3</td>
</tr>
<tr>
<td>Cones Approximation[29]</td>
<td>0.0</td>
</tr>
<tr>
<td>Cones Approx With Planar[29]</td>
<td>0.4</td>
</tr>
<tr>
<td>Cones Approx with Full Pose</td>
<td>0.9</td>
</tr>
<tr>
<td>Method of [58]</td>
<td>10.2</td>
</tr>
<tr>
<td>1 Image</td>
<td>6.32</td>
</tr>
<tr>
<td>3 Images</td>
<td>6.32</td>
</tr>
<tr>
<td>4 Images</td>
<td>3.95</td>
</tr>
<tr>
<td>6 Images</td>
<td>5.00</td>
</tr>
</tbody>
</table>

4.4 Experiment using the Caltech Face Database

To further test the performance of the face recognition technique based on the image registration algorithm, a second database of images that consisted of images acquired in an uncontrolled environment were used to evaluate the face recognition technique. These im-
4.4 Experiment using the Caltech Face Database

Figure 4.2: Sample images from the Caltech database. Both images contain the same subject, but in different locations, relative position of the camera, ambient lighting and expression.

ages, compared to the previous section using the Yale Face Database B, are more natural in that they do not have set locations for the camera and strobe to ensure consistent pose and illumination conditions. Also the scale between various images is slightly different. This experiment more closely resembles a “real world” application of using casual portrait type images.

4.4.1 Caltech Face Database

The database used for this experiment was the Caltech Face Database [42, 43]. This database consists for 27 different subjects for a total of 450 images. These images have differences in background, scale, and expression. There are some slight differences in pose and illumination, however these are mainly due to the changes in environment for the images and are still mostly of a frontal pose and bright illumination. An example of the images from this database are shown in Figure 4.2, where the images in 4.2a and 4.2b are both of the same subject, however the background, position, scale and expression is changed.

An experiment was conducted similar to the previous section, using the proposed algorithm implemented in MATLAB. Only the subjects in the database that consisted of more then 20 images were used as test cases, which was 19 subjects out the 27 that are present in
4.4 Experiment using the Caltech Face Database

Figure 4.3: Example of the type of images that were removed from the database due to poor quality. Figure 4.3a demonstrates underexposure while Figure 4.3b represents too drastic a change in scale for this given experiment.

the database. Upon review of the database, it was noted that some images in the database were found to be poor quality (underexposed images, significant scale change), an example of this is shown in Figure 4.3 and thus those images were removed from the database. This resulted in a removal of about 15 images total. A cropped image of only the subject face was used as the training image and only one image was used in the training step for each comparison. The evaluation of the performance of the comparison used the same criteria as the previous section. The scale and error parameters for the feature extraction and post-processing step respectively remained the same as the previous section.

4.4.2 Performance of the Proposed Method on the Caltech Database

The results from the experiment are listed in Table 4.11\textsuperscript{12}. These results indicate that the proposed method leads to a very low error rate and also very low in both the false rejection rate and especially false acceptance rate.

Of note, the images for subject 6, the majority of those images were blurred, and from

\textsuperscript{1}Subject 13 was removed due to the majority of images being underexposed.

\textsuperscript{2}Subject 14 was removed due to the majority of the images being underexposed and also consisting of large scale differences.
Table 4.11: Results from the experiment using the Caltech Face Database.

<table>
<thead>
<tr>
<th>Subject</th>
<th>FR$_{total}$</th>
<th>FA$_{total}$</th>
<th>Error Rate</th>
<th>FR$_{subject}$</th>
<th>FA$_{subject}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>0.67</td>
<td>0.00</td>
<td>0.67</td>
<td>14.29</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 2</td>
<td>0.44</td>
<td>0.00</td>
<td>0.44</td>
<td>10.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 3</td>
<td>0.00</td>
<td>0.22</td>
<td>0.22</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Subject 4</td>
<td>1.33</td>
<td>0.22</td>
<td>1.56</td>
<td>30.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Subject 5</td>
<td>2.22</td>
<td>0.00</td>
<td>2.22</td>
<td>43.48</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 6</td>
<td>0.22</td>
<td>0.22</td>
<td>0.44</td>
<td>5.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Subject 7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 8</td>
<td>1.78</td>
<td>0.00</td>
<td>1.78</td>
<td>40.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 9</td>
<td>0.44</td>
<td>0.67</td>
<td>1.11</td>
<td>9.52</td>
<td>0.70</td>
</tr>
<tr>
<td>Subject 10</td>
<td>0.67</td>
<td>0.00</td>
<td>0.67</td>
<td>12.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 11</td>
<td>0.22</td>
<td>0.00</td>
<td>0.22</td>
<td>4.55</td>
<td>0.00</td>
</tr>
<tr>
<td>Subject 12</td>
<td>0.22</td>
<td>0.67</td>
<td>0.89</td>
<td>5.26</td>
<td>0.70</td>
</tr>
<tr>
<td>Subject 13</td>
<td>0.22</td>
<td>0.44</td>
<td>0.67</td>
<td>5.00</td>
<td>0.47</td>
</tr>
<tr>
<td>Subject 14</td>
<td>1.78</td>
<td>0.22</td>
<td>2.00</td>
<td>40.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Subject 15</td>
<td>0.00</td>
<td>0.22</td>
<td>0.22</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Subject 16</td>
<td>0.22</td>
<td>0.44</td>
<td>0.67</td>
<td>5.00</td>
<td>0.47</td>
</tr>
<tr>
<td>Subject 17</td>
<td>0.33</td>
<td>0.22</td>
<td>1.56</td>
<td>27.27</td>
<td>0.23</td>
</tr>
<tr>
<td>Average</td>
<td>0.69</td>
<td>0.21</td>
<td>0.91</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
the results, the algorithm was still able to successfully maintain a low false rejection rate which is consistent with the image registration algorithms performance to distortions. The speed of execution was also evaluated and found that the experiment operated at just under 2 seconds per comparison.

While there is no external comparison of the results for the uncontrolled database used here, these results are consistent with that of the previous experiment using the controlled database. The results between the two databases compare quite similar to each other with the uncontrolled database having slightly better results. This could be explained by the fact that the controlled database has a wider difference in illumination and pose angles. The closeness in these results between the two databases and experiments validates the approach of image registration for face recognition.

4.5 Summary

This chapter has shown the performance of the proposed face recognition method using image registration. Using two separate applications of the face recognition problem with controlled and uncontrolled images the performance could be measured and compared against existing techniques. The results presented here show that the proposed method performs very well for each set of database images and has very good performance with respect to existing methods. The overall error rate using these two databases is very low and further shows that the image registration technique is a good method for face recognition. The addition of using multiple images in the training set provided improved results, specifically in the reduction of the false rejection rate. The cost of using multiple images in computation time of the face recognition algorithm is shown to be small.
Chapter 5

Conclusion and Future Work

5.1 Conclusions

The focus of this thesis has been the creation of a face recognition algorithm that is based on techniques from image registration. In Chapter 2, the concept of image registration was introduced. Specifically, a method of using scale-interactions of Mexican-hat wavelets and the magnitudes of the Zernike moments of image feature points was discussed. In Chapter 3, a face recognition algorithm using this image registration method was introduced. In Chapter 4, a group of experiments were conducted using databases acquired in controlled and uncontrolled environments to test the validity of the proposed face recognition algorithm.

5.1.1 Image Registration

In Chapter 2, an image registration method is discussed. The specific technique proposed was developed in [16, 17, 44] and was adapted here with improvements in the choice of scale parameters and parts of it were implemented in C to increase the computational speed. The algorithm itself is explained using four steps. First the feature points of the images to be registered are extracted, using the scale interactions of Mexican-hat wavelets for filtering the image while allowing for compensation in scale differences. Next, the Zernike
moments are calculated for each of the features points which are then used for determining the correspondence between the images. This correspondence is based on the similarity of the magnitude of the Zernike moments between the feature points of two images. Finally, an iterative recursive least-squares minimization is performed in order to remove any outlier feature point pairs, as well as to determine a set of affine parameters that can be used to transform one image onto the other.

The effect that scale differences between the two images to be registered has on the effectiveness of the image registration technique, specifically in the feature point extraction step, was discussed. The ratio between the scale parameters was more important than the value of the scale parameters. A method to efficiently determine the scale parameters and to provide a comprehensive range of scale ratios to be used for the feature extraction was discussed.

Examples illustrating the image registration technique were presented that included images of partial overlap, difference in scale, geometric distortions and degradations. Through these examples, it was shown that the image registration technique which was adapted from [16, 17, 44] provides excellent registration accuracy with several different image types.

5.1.2 Face Recognition

In Chapter 3, a face recognition algorithm is proposed that is based on the image registration technique from chapter 2. The proposed face recognition algorithm used the advantages that the image registration technique provided, in particular, with respect to using images of partial overlap, geometric distortions and changes in image scale. The face recognition algorithm was presented in three stages: a training step is performed on a set of training images consisting of cropped the facial images, for which the feature point extraction on these images is performed and the magnitudes of the Zernike moments are calculated. This information is then stored into a data file for later use. The second stage was to execute the feature extraction and registration algorithm provided by the image registration technique. Only the sample image is required to have the feature extraction step
performed which results in faster execution of the algorithm. Comparing the sample image data to the training images data will result in match or rejection. This decision is performed by the final step of the algorithm, the post-processing step. This post-processing step obtains three error parameters from the estimated transformation parameters found in the registration step and uses these parameters to decide if a match has been found.

Examples were performed that demonstrated the proposed face recognition algorithm against changes in pose, illumination, background and scale conditions. In each example, a demonstration with the same subject and a different subject was used to illustrate how the face recognition algorithm is able to successfully determine between matches and non-matches.

### 5.1.3 Experimental Results

In Chapter 4, two groups of experiments were performed that represented examples of face recognition applications. The first set of experiments made use of a database that was designed to provide a controlled set of high-resolution images of various subjects under differing conditions of pose and illumination. The second experiment used high-resolution images that were acquired in a largely uncontrolled environment. These images had differences in background, scale, expression, pose and illumination conditions.

In the first set of experiments, the Yale Face Database B was specifically used [29]. This database was structured to test against the effect that changing illumination and pose had on the face recognition algorithm accuracy. The results from these experiments showed that the face recognition algorithm from chapter 3 was very accurate in determining between acceptance and rejections of the sample images. Further, the addition of the multiple training images reduced the effect that pose and illumination conditions have on the accuracy of the face recognition algorithm. Comparing the results to other methods that made use of this database showed that the proposed method performed very well compared to existing techniques.

The second experiment group made use of the Caltech Face Database [42, 43]. This
database contained a large number of subject images that were acquired at different locations and under changing scale, pose, and illumination conditions. The results from this experiment show that the face recognition algorithm has very good accuracy for these type of images.

5.2 Future Work

The following sections provide some research topics that would extend the work on the face recognition algorithm presented in this thesis:

Currently in the feature extraction and registration step of the face recognition algorithm, a test set of images is required to determine the best values for the scale parameter for the feature point extraction. An improvement to this algorithm would be in determining the scale parameters for the feature point extractor automatically.

Considerations that would need to be taken into account for this type of feature would be the effect on the training step of the face recognition it would have. Currently the training step is pre-calculated for the feature point extraction, with the results found through the learning set. If this is removed from the learning set, then the training images would need to be pre-calculated using many more scale values, to ensure that a large number of scale ratios between the sample image are available.

A method that could be used to provide this automatic detection of the scale parameters would be to include a starting scale parameter and increase by an adjustable step increase. This step increase could be related to the size of the image to ensure that the scales are not increased too dramatically. The feature extraction would then iterate using this new scale parameter until only a few feature points are detected, meaning that the scale parameter has reached an upper limit.

A further application using the proposed face recognition algorithm is in the use of images which have multiple faces located in the image. This application can be best dealt with by combining object recognition techniques to first detect if a face a present on an
image and then use this information to crop the face from the image to be used by the face recognition technique presented in this thesis.
Bibliography


[18] H.-K. Kim, J.-D. Kim, D.-G. Sim, and D.-I. Oh, “A modified zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image re-


