# New Methods for Image Registration and Normalization using Image Feature Points

by

**Mohamed Seddeik Yasein**

M.A.Sc, University of Victoria, 2002
B.Sc, Suez Canal University, 1996

A Dissertation Submitted in Partial Fullfillment of the
Requirements for the Degree of

**Doctor of Philosophy**

in the Department of Electrical and Computer Engineering

# New Methods for Image Registration and Normalization using Image Feature Points

by

**Mohamed Seddeik Yasein**

M.A.Sc, University of Victoria, 2002
B.Sc, Suez Canal University, 1996

**Supervisory Committee**

---

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

---

Dr. Kin Fun Li, Department Member
(Department of Electrical and Computer Engineering)

---

Dr. S. Neville, Department Member
(Department of Electrical and Computer Engineering)

---

Dr. Sadik Dost, Outside Member
(Department of Mechanical Engineering)

# Supervisory Committee

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

Dr. Kin Fun Li, Department Member
(Department of Electrical and Computer Engineering)

Dr. S. Neville, Department Member
(Department of Electrical and Computer Engineering)

Dr. Sadik Dost, Outside Member
(Department of Mechanical Engineering)

# Abstract

In this dissertation, the development and performance evaluation of new techniques for image registration and image geometric normalization, which are based on feature points extracted from images are investigated.

A feature point extraction method based on scale-interaction of Mexican-hat wavelets is proposed. This feature point extractor can handle images of different scales by using a range of scaling factors for the Mexican-hat wavelet leading to feature points for different scaling factors. Experimental results show that the extracted feature points are invariant to image rotation and translation, and are robust to image degradations such as blurring, noise contamination, brightness change, etc. Further, the proposed feature extractor can handle images with scale change efficiently.

A new algorithm is proposed for registration of geometrically distorted images, which may have partial overlap and may have undergone additional degradations. The global $2D$ affine transformations are considered in the registration process. Three main steps constitute the algorithm: extracting feature point using a feature point extractor based on scale-interaction of Mexican-hat wavelets, obtaining the correspondence between the feature points of the reference and the target images using Zernike moments of neighborhoods centered on the feature points, and estimating the transformation parameters between the first and the second images using an iterative weighted least squares algorithm. Experimental results show that the proposed algorithm leads to excellent registration accuracy using several types of images, even in cases with partial overlap between images. Further, it is robust against many image degradations and it can handle images of different scales effectively.

A new technique for image geometric normalization is proposed. The locations of a set of feature points, extracted from the image, are used to obtain the normalization parameters needed to normalize the image. The geometric distortions considered in the proposed normalization technique include translation, rotation, and scaling. Experimental results show that the proposed technique yields good normalization accuracy and it is robust to many image degradations such as image compression, brightness change, noise contamination and image cropping.

A blind watermarking technique for images is proposed, as an example of the possible applications of the presented geometric normalization technique. In order to enhance robustness of the watermarking technique to geometric distortions, the normalization technique is used to normalize the image, to be watermarked, during the embedding process. In the watermark detection stage, the normalization parameters for the possibly distorted watermarked image are obtained and used to transform the watermark into its normalized form. The transformed watermark is, then, correlated with the image to indicate whether the watermark is present in the image or not.

Experimental results show that the proposed watermarking technique achieves good robustness to geometric distortions that include image translation, rotation, and scaling.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

| | |
|---|---|
| 2D | Two-Dimensional |
| dB | decibel |
| FFT | Fast Fourier Transform |
| G. N. | Geometric Normalization |
| IFFT | Inverse Fast Fourier Transform |
| JPEG | Joint Photographic Experts Group |
| LoG | Laplacian of Gaussian |
| PSNR | Peak Signal to Noise Ratio |
| RANSAC | RANdom SAmple Consensus |
| RMSE | Root Mean Squared Error |
| SIFT | Scale Invariant Feature Transform |
| SPOT | Satellite Pour l'Observation de la Terre |
| STD | STandard Deviation |
| SUSAN | Smallest Univalue Segment Assimilating Nucleus |
| TM | Thematic Mapper |

# Acknowledgment

In the name of Allah, the most Gracious, the most Merciful. All praise be to Allah the Almighty who has given me knowledge, patience, and perseverance to finish my Ph.D. dissertation.

My great thanks to my parents who stood all the way behind me with their support, encouragement, and prayers until this work was done. I would also like to express my special thanks to my wife for her advice and support.

My deepest thanks to my supervisor Dr. Panajotis Agathoklis for his invaluable scholarly advice, inspirations, help, and guidance that helped me through my Ph.D. dissertation work. I will always be indebted to him for all he has done for me, and it is a pleasure to acknowledge his guidance and support. Thank you very much for being such a fantastic supervisor.

I would like to acknowledge the advice and support from my supervisory committee members: Dr. Kin Fun Li, Dr. Stephen W. Neville, and Dr. Sadik Dost for making my dissertation complete and resourceful.

I feel a special gratitude to Dr. M. Watheq El-Kharashi for his guidance, support, and beneficial discussions. He has provided me with so much help and valuable advices.

Finally, I would like to thank my friends Ahmed Awad, Mohamed Fayed, Omar Hamdy, Yousry Abdel-Hamid, Haytham Azmi, Ahmed Morgan, Mohamed El-Gamal, Emad Shihab, Khaled Khayyat, and Abdelsalam Amer for their generous friendship and enlightening discussion.

*M. S. Yasein, Victoria, BC, Canada*

# Dedication

*To my parents,*

*my wife,*

*and my sons.*

# Chapter 1

# Introduction

A key issue in many application areas such as computer vision, remote sensing, medical imaging, pattern recognition and image retrieval is to deal with images that are acquired by imperfect imaging systems. Different image acquisition conditions, such as imaging geometry, sensing/environmental conditions and sensor errors, may introduce several types of distortions and degradations to the observed image, compared to the original scene (image). This may result in images that suffer degradations such as contamination with noise, blurring and brightness/contrast change. Moreover, image-planes may have different orientations, scales, positions or may have undergone geometric distortions.

One of the important applications that concerns geometric distortions of images is image registration. Given two, or more, images to be registered, an image registration system estimates the parameters of the geometric transformation model that maps a given target image, that was taken from different viewing positions or at different times, to the reference one. Registration methods are increasingly in demand by many image analysis and processing systems, as the first step, to accurately capture the geometric transformations of the image data. For example, image registration is used in image stitching, where multiple images are combined to produce a panorama

or larger image. Image registration is used in the analysis of remotely sensed data where an image must be transformed, using image registration techniques, to match the orientation and the scale of previously acquired images. Also, in motion analysis applications, image registration is utilized to find changes between two subsequent frames in a video sequence and in medical imaging systems, images need to be aligned first which can be accomplished through registration process, before they can be compared for analysis and diagnostic purposes.

Another application that concerns geometric distortions of images is image geometric normalization, which has been used as a preprocessing step in many applications such as pattern recognition, image retrieval and image watermarking. Image normalization is used in such applications with the purpose of representing objects, patterns (or the entire image) regardless of changes in their orientation, size, or position [1]. The use of image normalization as a preprocessing step might limit the range of variations of images/patterns, as it effectively decouples the problem of image deformations from the main task of the application, for example, retrieval, recognition or classification, etc [2].

## 1.1 Image Registration

Image registration has found applications in numerous real-life applications such as remote sensing, medical image analysis, computer vision and pattern recognition [3]. Given two, or more, images (views) to be registered, image registration estimates the parameters of the geometric transformation model $\tau_{G_R}$ that maps a given target image to the reference one. The transformation model $\tau_{G_R}$ maps a point $\mathbf{x} = (x_1, x_2)$ of the reference image $I$ to a point $\mathbf{x}^{'} = (x_1^{'}, x_2^{'})$ of the target image $I^{'}$ as $\mathbf{x}^{'} = \tau_{G_R}(\mathbf{x})$. If a pixel $\mathbf{p}$ in $I$ corresponds to $\mathbf{p}^{'}$ in $I^{'}$, then an accurate registration should result in $\tau_{G_R}(\mathbf{p})$ as close as possible to $\mathbf{p}^{'}$. A registration error is, therefore, the result of

the displacement $\tau_{G_R}(\mathbf{p}) - \mathbf{p}'$. The basic idea of registration systems is demonstrated in Fig. 1.1


Figure 1.1: Image registration concept.

## 1.2 Image Geometric Normalization

The basic idea of image geometric normalization is to transform a given image into a standard form where this normalized form is independent of any possible geometric distortions applied on the image [4]. Given an input image, geometric normalization systems designed to geometrically transform this image into a standard form such that the normalized image is invariant to geometric distortions. In addition, a robust normalization system should be able to perfectly normalize images, regardless of any additional image degradation such as noise contamination, cropping, etc. The basic idea of robust geometric normalization systems is demonstrated in Fig. 1.2. In general, the geometric distortions considered in geometric normalization systems are rotation, scaling and translation.

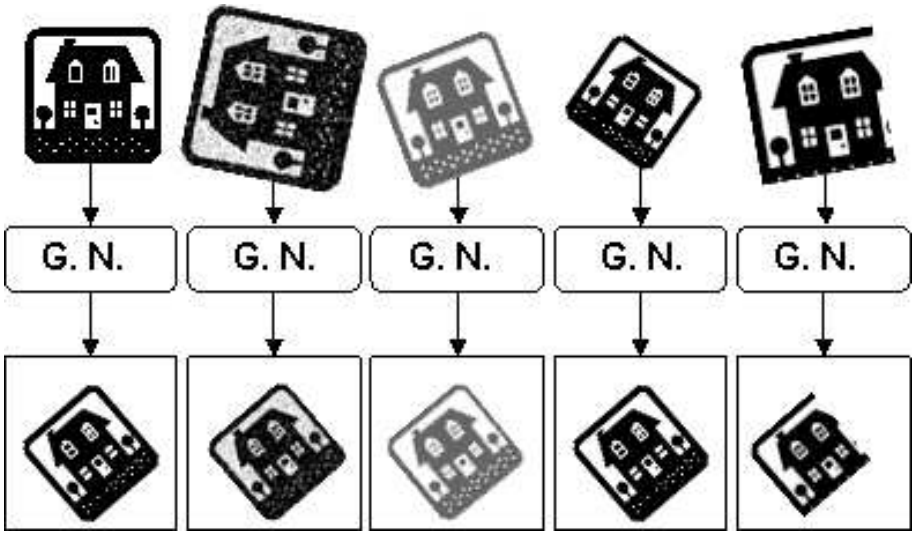



Figure 1.2: Image Geometric Normalization (G. N.): The first row contains the input images to be normalized and the third row contains the corresponding resulting normalized images.

An effective approach that can be utilized in applications that involve image registration or image normalization is through using feature point extraction. Image feature point extraction methods have been an area of interest to researchers in image processing due to their potential applications to several image analysis problems.

In image registration techniques, feature point extraction can be utilized to extract two sets of feature points from the reference and the target images. Next, the correspondence between the feature points of the two images is established. Finally, the geometric transformation parameters between the two images are estimated.

In image normalization techniques, feature point extraction can be utilized to extract the feature points from the given image and then the normalization parameters are estimated based on the geometric properties of image feature points.

These image registration and normalization techniques are usually designed with the objective to be robust to different types of image distortions and degradation.

Through out this dissertation, image distortions will refer to geometric transformations such as rotation and scaling while image degradation will refer to the reduction in the image quality caused by, for example, noise contamination, blurring and brightness/contrast change.

## 1.3 Scope and Contributions of the Dissertation

This dissertation is organized as follows.

In Chapter 2, some basic concepts of image feature point extraction, image registration and normalization are presented to provide a basis for the subsequent chapters. A discussion and literature review of the most relevant concepts, terminologies and techniques pertaining to image registration and image normalization are presented.

In Chapter 3, a feature point extractor based on scale-interaction of Mexican-hat wavelets to extract a set of feature points from an image is presented. This feature point extraction method is an extension of an earlier technique allowing its effective use with images of different scales. Its performance is illustrated with experiments and compared with other techniques. This feature point extractor will be used in the image registration and image normalization techniques to be presented in Chapter 4 and in Chapter 5, respectively.

In Chapter 4, a new algorithm for image registration is presented. The objective of this algorithm is to register images that are geometrically distorted and in addition may have partial overlap and/or they may have undergone degradations caused by noise contamination, blurring, etc. The geometric distortions considered in the registration process are the global $2D$ affine transformations. Experimental results and comparison with existing registration techniques illustrate the accuracy of image registration for various types of such as aerial images, digital photography, etc.

In Chapter 5, a technique for image normalization based on the geometric properties of image feature points is proposed. The proposed technique can be used to normalize input images independent of possible distortions such as rotation, translation and scaling. Experimental results illustrate the accuracy of normalization and the robustness against several geometric distortions, image degradations and common image-processing operations. A technique for image watermarking with robustness to geometric distortions is introduced as an example of the possible applications of the proposed geometric normalization technique. Performance evaluations show that the proposed watermarking technique achieves good robustness to geometric distortions that include image translation, rotation, and scaling.

In Chapter 6, the results and contributions of this dissertation are summarized and directions for future research are suggested.

# Chapter 2

# Introduction to Image Geometric Transformations and Feature Point Extraction for Image Registration and Normalization Systems

In this chapter, the fundamental aspects of geometric transformations of images and some of the related applications are presented to provide a basis on which the subsequent chapters are based. The chapter begins with some background knowledge about different types of geometric transformations and then, in Section 2.2, image feature point extraction techniques are discussed. Section 2.3 and Section 2.4 discuss some image applications that are rely on feature point extraction; these applications include image registration and image normalization. Finally, Section 2.5 summarizes the chapter.

## 2.1 Images and Geometric Distortions

A geometrical transformation $\tau_G$ is a mappings of points (pixels) from an image to the transformed points of a second image. The transformation $\tau_G$ applied to a point $\mathbf{x} = (x_1, x_2)$ of an image $I$ produces a transformed point $\mathbf{x}^{'} = (x_1^{'}, x_2^{'})$ of the transformed image $I^{'}$ such that $\mathbf{x}^{'} = \tau_G(\mathbf{x})$ [5].

The most common global geometric distortions are the rigid geometrical transformations. These transformations preserve all distances and also preserve the straightness of lines. In addition, the overall geometric relationships between points do not change and, in consequence, the shapes of objects in the image do not change. In this type of transformations, there are two components to the specification, a translation and a rotation components. The translation is a two-dimensional vector that may be specified by giving its two parameter in $x$ and $y$ directions while the rotation angle can be specified by one parameter. Hence, a combined transformation of these types typically has tree parameters, $t_{x1}$, $t_{x2}$ and $\theta$, which maps a point $(x_1, x_2)$ of the first image $I$ to a point $(x_1^{'}, x_2^{'})$ of the transformed image $I^{'}$ as follows:

$$I^{'} = \tau_G(I) = T(t_{x1}, t_{x2}) + R(\theta)I, \tag{2.1}$$

where $T$ and $R$ represent translation and rotation operations respectively. In detail, this can be represented as

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix} = \begin{pmatrix} t_{x1} \\ t_{x2} \end{pmatrix} + \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{2.2}$$

where $t_{x1}$ and $t_{x2}$ are translation parameters in $x$ and $y$ directions respectively and $\theta$ is a rotation angle.

The aforementioned transformations can be accompanied with uniform image scaling and, therefore, a combined transformation of this type has four parameters,

$t_{x1}$, $t_{x2}$, $s$ and $\theta$, which can be defined as

$$I^{'} = \tau_G(I) = T(t_{x1}, t_{x2}) + sR(\theta)I \tag{2.3}$$

In detail, this can be represented as

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix} = \begin{pmatrix} t_{x1} \\ t_{x2} \end{pmatrix} + s \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{2.4}$$

where $s$ is the scaling parameter. Examples of rigid transformations are shown in Fig. 2.1.

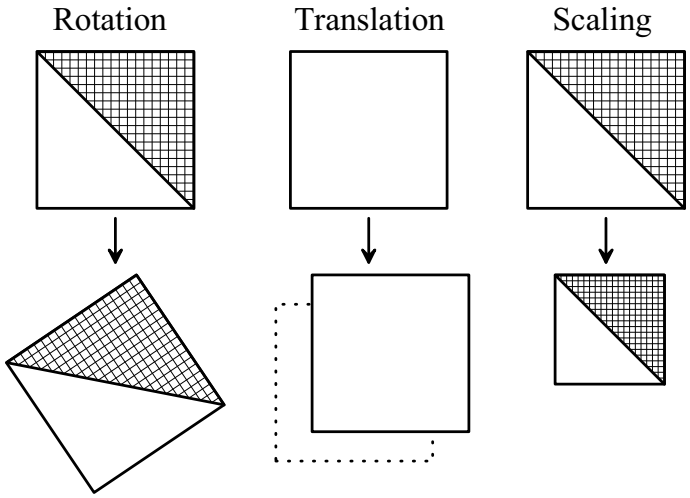



Figure 2.1: Examples of rigid transformations.

Fig. 2.2 shows examples of images that demonstrate the aforementioned geometric transformations.

A more general type of rigid global transformations is the $2D$ affine transformation. Examples of affine transformation are image shearing, in $x$ or $y$ directions, and changes in aspect ratio due to nonuniform scaling. The general form of affine transformation can be represented as

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} t_{x1} \\ t_{x2} \end{pmatrix}, \tag{2.5}$$

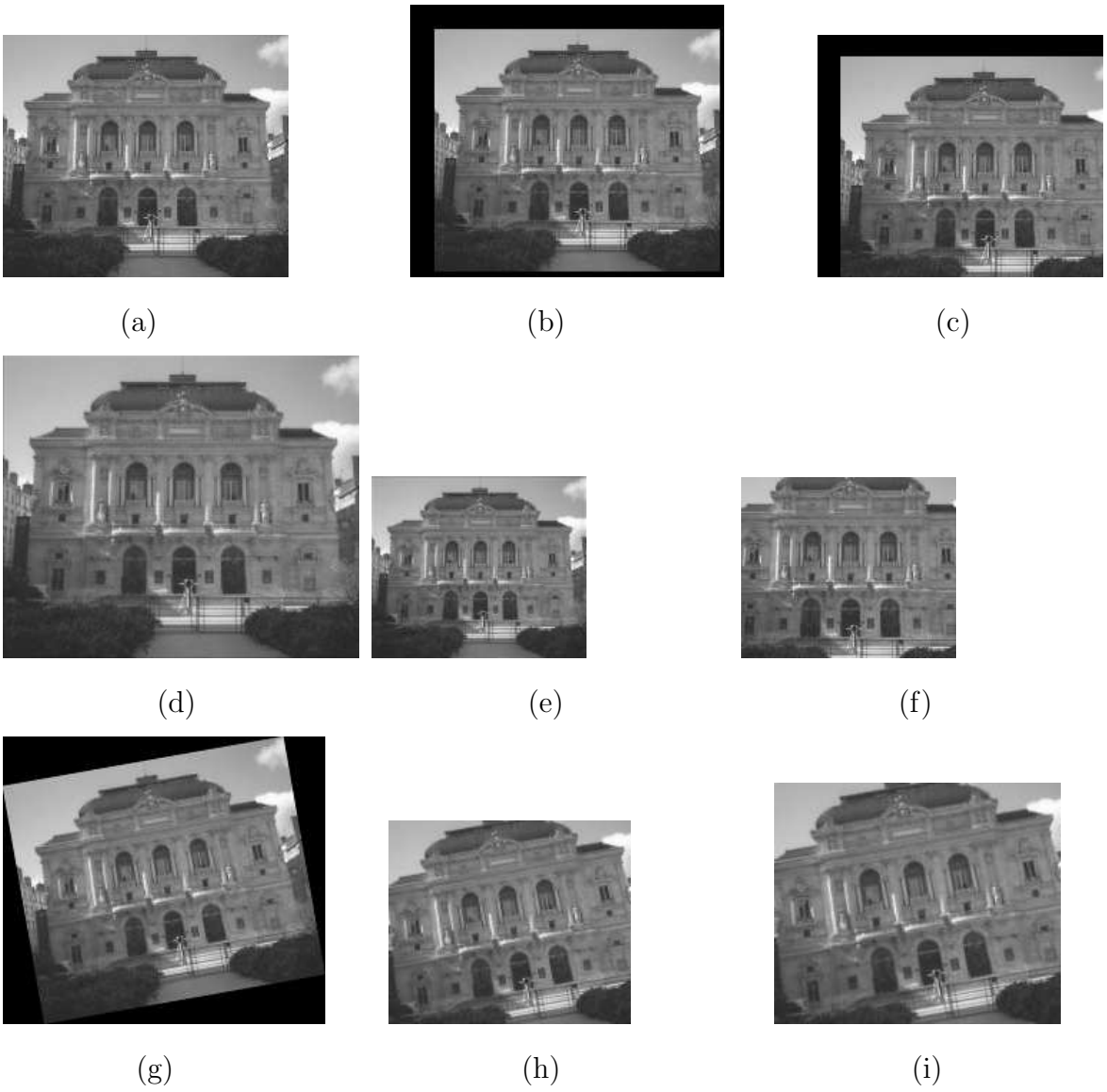



Figure 2.2: An image and the results of applying different geometric operations: (a) Original image, (b) Translation without cropping, (c) Translation with cropping, (d) Scaling up by 150%, (e) Scaling down by 75%, (f) Central cropping of 25%, (g) Rotation by $10^o$, (h) Rotation and cropping by $10^o$, (i) Rotation and rescaling by $10^o$.

where $a_{11}$, $a_{12}$, $a_{21}$, $a_{22}$, $t_{x1}$, and $t_{x2}$ are the transformation parameters.

Projective transformation is another type of transformation that is obtained when adding two more parameters to the above transformation and that introduces an additional distortion in the image [6,7]. This transformation describes what happens when viewing an object from some arbitrary viewpoint at a finite distance. It maps lines to lines, but does not necessarily preserve parallelism. The general form of Projective transformations can be represented as

$$\begin{aligned} x_1^{'} &= \frac{a_{11}x_1 + a_{12}x_2 + a_{13}}{a_{31}x_1 + a_{32}x_2 + 1} \\ x_2^{'} &= \frac{a_{21}x_1 + a_{22}x_2 + a_{23}}{a_{31}x_1 + a_{32}x_2 + 1}, \end{aligned} \tag{2.6}$$

where $a_{11}$, $a_{12}$, $a_{13}$, $a_{21}$, $a_{22}$, $a_{23}$, $a_{31}$, $a_{32}$ and $a_{33}$ are the transformation parameters. Examples of affine and projective transformations are shown in Fig. 2.3.

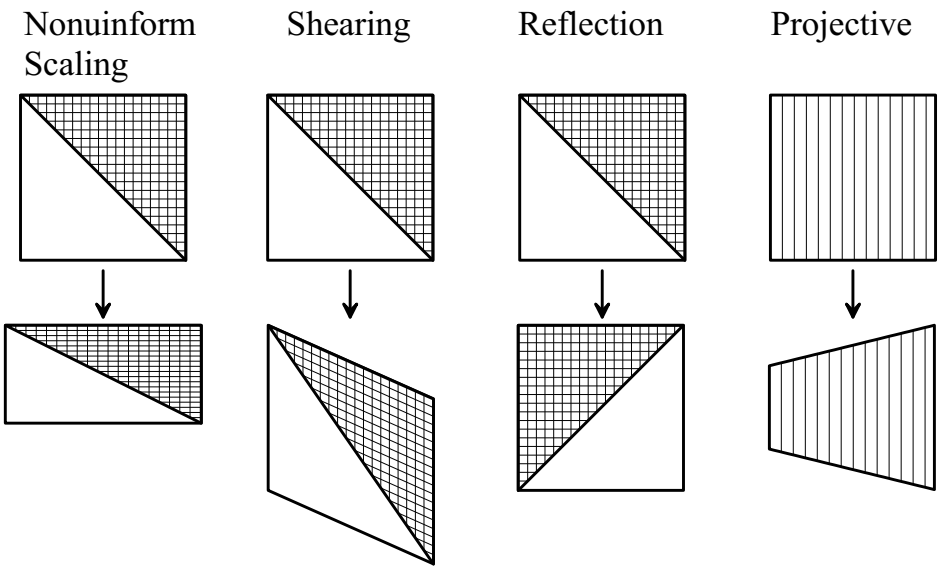



Figure 2.3: Examples of affine and projective transformations.

Another geometric distortions, which can occur, include bilinear and curved transformation as in 2.7 and 2.8, respectively. The general forms of these trans-

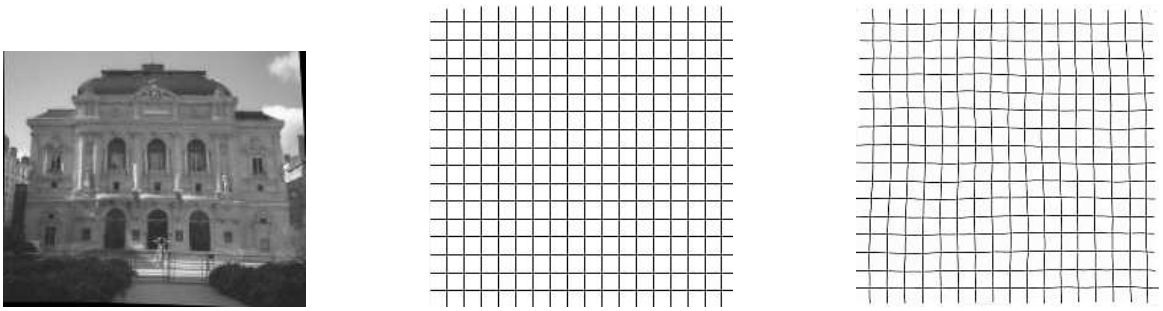
formations can be represented as

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix} = x_1 x_2 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} t_{x1} \\ t_{x2} \end{pmatrix} \tag{2.7}$$

and

$$\begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} ((1-\beta)a_{11} + \beta a_{12}) \sin(\alpha\pi) \\ ((1-\beta)a_{21} + \beta a_{22}) \sin(\alpha\pi) \end{pmatrix} \tag{2.8}$$

A more complex distortion includes applying the aforementioned transformation locally in small blocks instead of the whole image area [8]. The visual distortion due to this kind of local geometrical transformation are more difficult to model, even with a simple local transformation like the one in 2.4. Fig. 2.4-a shows the image in Fig. 2.2-a after applying local random distortions. As shown in the figure, the effect of such local transformation might be unnoticeable. In order to demonstrate the effect of the transformation, a grid image in Fig.2.4-b is operated by the transformation; the result is shown in Fig.2.4-c after applying local random distortions.



(a) (b) (c)

Figure 2.4: (a) A local random distorted image , (b) Original grid image , (c) Local random distorted grid image.

## 2.2 Feature Point Extraction from Images

Feature point and corner detectors have been an area of interest to researchers in image processing. There has been much effort to develop precise, robust and fast methods for feature point extraction. A point in an image is considered a feature point if it has the properties of distinctiveness and invariance [9]. Several approaches have been developed for feature point extraction and the points extracted by these methods differ in locations and structure, for example, edges, corners, blob-like shapes, etc. In general, the objective is to develop a feature point extractor that is robust to the most common geometric transformations and any possible degradation introduced by different viewing conditions.

In [10], a robust corner detector that is based on the first-order derivatives was developed. A better known corner detector is the 'Harris detector', which has been presented in [11]. It is based on the properties of the image gradients. Another intuitive approach, called SUSAN (Smallest Univalue Segment Assimilating Nucleus) method, was presented in [12]. The concept of each image point having a local area of similar brightness associated with it is the basis for the SUSAN principle. An example that illustrates the feature point extracted using Harris detector and SUSAN method is shown in Fig. 2.5. Recently, a parametric corner detector that does not employ any derivatives was designed to handle blurred and noisy data in [13]. In [14], an algorithm that uses optical flow ideas to extract the features in images was presented. Another method that is based on the wavelet transform was presented in [15]. The initial control points are obtained in the lowest resolution of the wavelet decomposition and then are refined at progressively higher resolutions. Some techniques use closed-boundary regions as the features such as the one in [16] where a two-threshold method is used to extract well defined contours in the images and the centers of gravity of such regions are taken as the feature points. A similar method was proposed in [17].

In this method, images are segmented by the improved Laplacian of Gaussian (LoG) edge extraction technique to define regions and the centers of gravity of such regions are taken as the feature points.

In recent years, there has been a lot of research in finding image representations over several scales. The concept of scale-space representation [18] has been developed by the computer vision community to deal with image structures at different scales. The scale-space representation turns out to be a solution to the diffusion equation with the image as the initial condition [18]. Convolving an image with Gaussian function of varying variance is equivalent to solving the linear heat diffusion equation with the image as the initial condition. The approach proposed in [19] combines a feature points detector and descriptors based on the gradient distribution in the detected regions. The resulting technique, the Scale Invariant Feature Transform (SIFT), has been used successfully in many computer vision applications. This approach is based on detecting feature points in the scale-space generated by the difference of Gaussians (DOG), which are invariant to rotation and scaling. A point's descriptor is obtained by computing a histogram of local oriented gradients around the feature points and stores the bins in a 128−dimensional vector. In [19,20] point detectors that are robust with respect to affine transformations and view angle changes and descriptors that can be used to perform reliable matching between different views of an object or scene for computer vision and pattern recognition applications have been presented.

## 2.3 Image Registration Techniques

Many image registration techniques have been proposed in the literature. In general, existing image registration techniques can be categorized into two classes [21]: area-based and feature-based methods. An extensive survey over image registration techniques can be found in [3, 21, 22]. Area-based methods register the images by

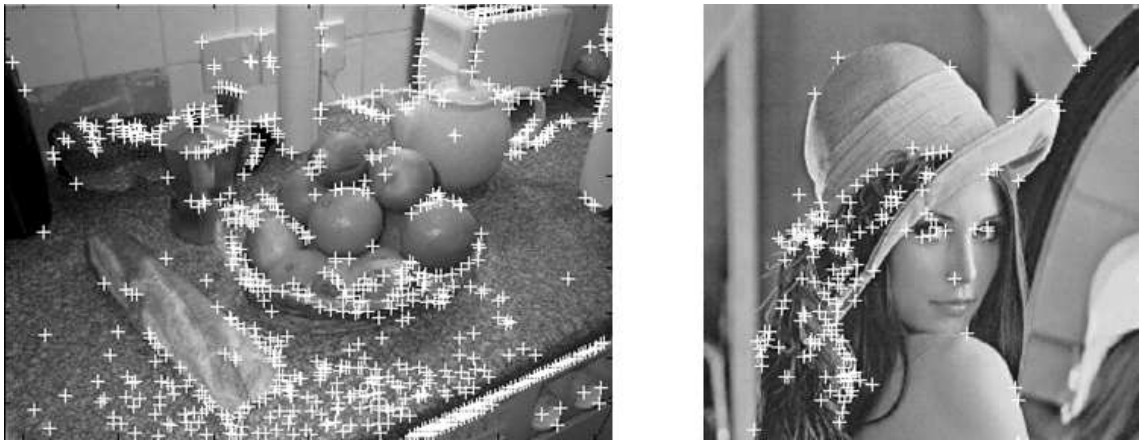


Figure 2.5: Feature point extraction from two images. The extracted points superimposed on each image, the left image using Harris detector and the right image using SUSAN method.

means that use the pixels intensities in image areas. Examples of this category include mutual information-based methods [23,24], Fourier transform-based methods [25–27], etc. On the other hand, feature based techniques attempt to extract features from images and use these features to obtain the transformation parameters for registering the images. Such image feature points may be based on some image characteristics such as corners, line segments, curves, object contours and feature points of the image, as in [16,28,29]. The performance of such techniques depends on several factors, such as the area of overlap between images and to what extent it is possible to model the different orientation between images with simple geometric transformations. Further, image quality, affected by degradations such as noise contamination and blurring, as well as, image characteristics such as smooth/textured areas or similarity of different areas, play also a role in the techniques' performance. A good overview can be found in [3, 21, 22]. In the following, some of the techniques that are relevant to the discussion in the subsequent chapters are briefly discussed.

In [16], a registration technique that uses closed-boundary regions as the features

is presented. In this technique, a two-threshold method is used to extract well-defined contours in images and the centers of gravity of such regions are taken as feature points. In the feature-matching step, chain code representation of contours was proposed as the invariant descriptors and a chain code correlation-like measure was used for finding the correspondence. Transformation parameters are estimated based on the final matched point pairs. A similar method was proposed in [17] where images are segmented by the improved Laplacian of Gaussian (LoG) edge extraction technique to define regions and the centers of gravity of such regions are taken as the feature points. Region correspondences are obtained using invariant moment shape descriptors with chain-code matching. A method that uses control points and moment invariants was proposed in [30, 31]. To find feature points in images, a parametric corner detector, developed for detection of corner-like dominant points in blurred images [13] was used. The matching is performed by means of a class of moment invariants, which are rotation, blur invariant, and are based on complex moments. In [32], an automated matching based on normalized cross correlation is used to generate a number of control points in the images. Then, robust estimation of the mapping function is performed using the random sample consensus (RANSAC) algorithm, which gives the robustness to this method. A system for image registration called imREG [1] was presented in [33]. In this tool, three different algorithms for registration have been implemented. The first algorithm uses optical flow idea [14] to extract features in both images. An initial corresponding set of control points in the two images is obtained and another refinement is performed by using a purely geometric matching procedure. The second algorithm is based on the wavelet transform [34]. The initial control points are obtained in the lowest resolution of the wavelet decomposition and then refined at progressively higher

[1] http://nayana.ece.ucsb.edu/registration/

resolutions. The point matching process uses the correlation coefficient as a similarity measure and a consistency-checking procedure to eliminate incorrect control points. The third method is based on contour matching. Chain code correlation and other shape similarity criteria such as moments are used to match closed contours and then a consistency check is conducted to eliminate false matches. The transformation function is estimated using the least squares method. Another approach, that does not require establishing explicit point correspondences, is proposed in [28]. This method is based on matching curves between the related images by superimposing two related curves upon each other and fitting it with a single B-spline. A technique proposed in [35] is based on computing geometric moment invariants for a window sliding across the whole image, considering translated and rotated images. A technique proposed in [36] is based on the use of Zernike moments and RANSAC robust fitting to guarantee stability, and Kanade-Lucas-Tomasi (KLT) tracker to provide accuracy. In [37], rather than using Zernike moments as descriptors, they assume that a set of points and their correspondents are given and the properties of Zernike moment are used to obtain a scaling and a rotation factor while the translation factor are obtained using the fast Fourier transform (FFT).

## 2.4 Image Geometric Normalization Techniques

Image Geometric normalization techniques have been developed as an elegant preprocessing method that transforms the given distorted input image into its normalized form such that it is invariant under geometric transformations [4]. Another class of image normalization techniques handles only image variations that are related to conditions of image acquisition, such as noise contamination and illumination/color change [38, 39]. In general, geometric normalization techniques work for normalizing input images in a rotation, translation and scale invariant manner. A close look at the

existing normalization techniques show that it is hard to find a universal technique that suit all types of image characteristics, such as, binary, gray or color images, or entire image/region-based methods. Further, in most applications, it's required that the normalization techniques be robust to any image degradation caused by image compression, filtering, or noise contamination.

Part of the focus in this dissertation is image geometric normalization. Therefore, through out the rest of this dissertation, the term "image normalization" will be used to refer to "image geometric normalization" for simplicity's sake.

Several image normalization methods have been proposed in the literature. Many methods are implemented based on image moment invariants [40, 41]. Moment invariants are functions of the image moments, which are constructed to be invariant to geometric transformations and are used to obtain the normalization parameters. The moment concept has been introduced to pattern recognition by [40]. Since then, varieties of new moment types and moment-based methods have been developed and used [42, 43]. Moments are attractive because their computation is simple and uniquely defined for any image function.

The existing classical methods that rely on image moments adjust the coordinate system in a first step by moving its origin into the image centroid, which is obtained using the zero and the first-order central moments. Expressions to determine orientation and scale exploit the particular changes that moments experience under rotations and scale changes of an image. There are many techniques developed for detecting a pattern's orientation, such as the ones based on principal axes in [44, 45] and on the tensor theory in [1]. As for scale, a simple method that relies on the pattern size, which is a function of the zero-order moments, can be used.

Examples of moment based normalization techniques can be found in [4, 46–49]. Most of these methods give excellent normalization results under the condition that there is no other image degradation besides the geometric distortions. To reduce noise

sensitivity of the moment-based image normalization methods, some modifications are proposed in [50].

A different approach for normalization is based on object outlines and other features. In [2, 51], a method for curve normalization with respect to affine transformations is proposed. Curves estimated from object contours are first modeled. Then the sampled curve is normalized in order to achieve translation, scaling, skew, starting point, rotation, and reflection invariance. The normalization is based on a combination of curve features including moments and Fourier descriptors.

Another method for silhouette normalization was presented in [52]. The objective of this method is to reduce the sensitivity of the localization, orientation and scale estimation with respect to silhouette deformations that may be caused by possible changes in the region of the periphery of the silhouette. In order to achieve that, a robust statistics technique, as well as a shape dependent weighting function, are utilized to make the estimations much less sensitive to deformations.

## 2.5 Conclusions

The background aspects of geometric transformations of images and some of the related issues that are necessary for the development of an efficient feature point extractor and new image registration and image normalization techniques in the following chapters have been reviewed. Several types of geometric transformations have been presented. In addition, related applications that include image registration and image normalization techniques have been discussed.

# Chapter 3

# Feature Point Extraction using Scale-Interaction of Mexican-hat Wavelets

In this chapter, a feature point extraction method based on Mexican-hat wavelets will be presented. This method will be used in image registration and normalization methods, discussed in the next chapters.

This chapter is organized as follows. Section 3.1 discusses feature point extraction process: its objectives and requirements. Section 3.2 presents a feature point extraction method that is based on scale-interaction of Mexican-hat wavelets. Section 3.3 summarizes the chapter.

## 3.1 Introduction

Feature point and corner detectors have been an area of interest to researchers in image processing. Different image features may be utilized in different applications. Moreover, for a given task, different feature sets may be suitable for different algorithms employed to perform the same task.

### 3.1.1 Objectives

This chapter presents a feature point extraction method, which is required to satisfy the following objectives:

- The extracted feature points should be invariant to rotation and translation. In addition, in some applications, it is required that moderate affine transformations should not destroy or alter the extracted features.
- The features should be robust to moderate variations in scale and the feature extractor should have the ability to be modified to adapt image structures at different scales.
- The extracted features should have a well-localized support in the image, i.e., cropping parts of the image should not alter the remaining feature points.
- The extracted features should have a reasonable robustness to possible degradations such as noise contamination, lossy compression, etc.
- The extracted features should be suitable for more than one type of tasks.

## 3.2 Feature Point Extraction using Mexican-hat Wavelets

A feature point extraction method that is based on scale-interaction of Mexican-hat wavelets is presented here. This method is based on finding the maxima of the response of a feature detection operation, which involves convolving the image with the Mexican-hat wavelets.

### 3.2.1 Background

Feature point extraction using scale interaction of Gabor wavelets was proposed in [53] based on a model of end-stopped simple cells developed in [54]. End-stopped cells are

neurons in the visual cortex that respond strongly to line-ends, corners and highly curved segments [54] and can be simply modeled by taking the difference of the responses of two receptive fields at the same position and orientation but of different size. A receptive field of a cell is the area of the visual field (or the area on the retina), which responds when stimulated by light. The receptive field profile is the sensitivity profile of the cell and can be modelled by a spatial filter function [54]. It has been mathematically described in several different ways such as difference of Gaussian, Gabor wavelets, and derivative of Gaussian. Thus, the cell response can be treated as a spatial convolution between an image and the receptive filed profile. The Mexican-hat is a Laplacian of a Gaussian and its isotropic property makes it insensitive to orientation and a good choice for feature extraction [55]. A Mexican-hat wavelet has the shape of a signal with a positive peak in a negative dish. Convolving an image with Mexican-hat wavelets results in a response, which more likely detects blob-like shapes, bright areas surrounded by dark pixels or vice versa. Varying the width of the central peak of the Mexican-hat wavelet controls the size and the shape of the response. Mexican-hat based feature extraction methods such as the ones in [56] and its further development in [29, 57] give good results when there is no scale change between the images. An illustration of this feature extraction process is shown in Fig. 3.1.

The method, which will be presented here, is an extension of our technique presented in [57] to accommodate the scale change of images in a consistent manner.

### 3.2.2 The Feature Point Extraction Process

The method presented in [57] is based on finding the maxima of the response of a feature detection operation that involves convolving the image with the Mexican-hat

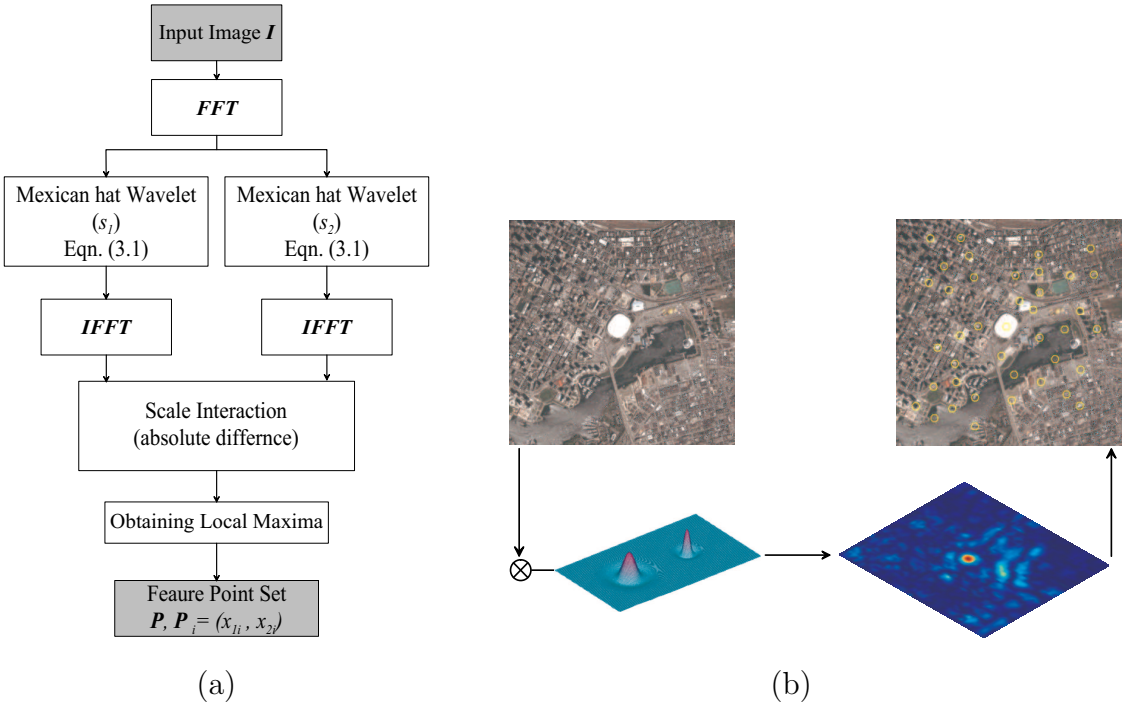



Figure 3.1: Feature point extraction: (a) Block diagram of the feature extraction process, and (b) Illustration of the feature extraction process showing the input image, Mexican-hat wavelets with two scales, the obtained response of applying Mexican-hat wavelets, and the extracted points superimposed on the image.

wavelet given by

$$Mex(\mathbf{x}, s_m) = \frac{1}{\sigma}\left(2 - \frac{x_1^2 + x_2^2}{\sigma^2}\right) e^{\left(-\frac{1}{2}\frac{x_1^2+x_2^2}{\sigma^2}\right)}, \tag{3.1}$$

where $\sigma = 2^{-s_m}$, $s_m$ is the scale of the function, $x_1$ and $x_2$ are the vertical and horizontal coordinates respectively.

This technique gives good results when there is no scale change. Scaling the images (performing a zooming operation, for example) has the effect that the relationship between the size of the center peak of the Mexican-hat and corresponding image areas is different. This results in extracting image feature points, which are not corresponding to the same locations in the images. To compensate for the scale difference, the Mexican-hat can be resized using a scaling factor $s_{p_i}$ as follows:

$$Mex(\mathbf{x}, s_m, s_{p_i}) = Mex(\frac{\mathbf{x}}{s_{p_i}}, s_m) = \frac{1}{\sigma}\left(2 - \frac{(\frac{x_1}{s_{p_i}})^2 + (\frac{x_2}{s_{p_i}})^2}{\sigma^2}\right) e^{\left(-\frac{1}{2}\frac{(\frac{x_1}{s_{p_i}})^2+(\frac{x_2}{s_{p_i}})^2}{\sigma^2}\right)}, \tag{3.2}$$

This scaling factor compensates for the unknown scaling difference between the two images. The feature extraction operation is applied using a range of scaling factors $s_{p_i}$ for the Mexican-hat wavelet and feature points for different scaling factors are extracted. The effect of this approach will be illustrated with an example at the end of this section after the proposed technique is presented.

The Mexican-hat wavelets-based feature extraction method involves two stages: finding the response of the image to a feature detection operation and localizing the feature points by finding the local maxima in the response. The feature extraction process is shown in Fig. 3.2.

1. The first stage finds the response $\phi(\mathbf{x}, s_1, s_2, s_{p_i})$ of the image to a feature detection operation as

$$\phi(\mathbf{x}, s_1, s_2, s_{p_i}) = \left|\Re(\mathbf{x}, s_1, s_{p_i}) - \Re(\mathbf{x}, s_2, s_{p_i})\right|, \tag{3.3}$$

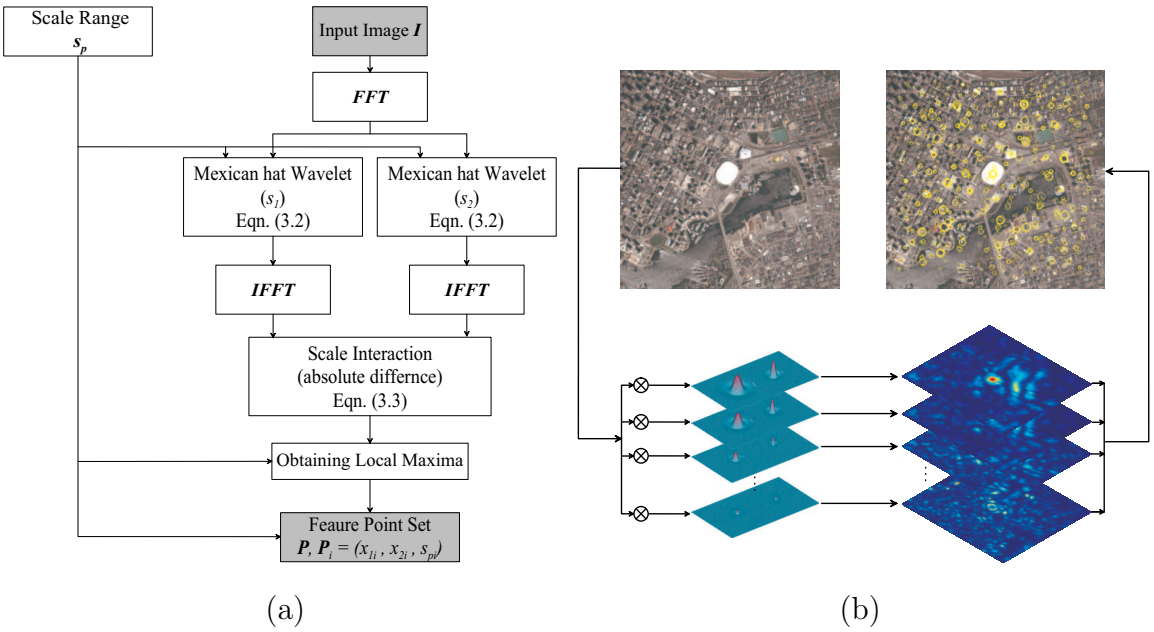



Figure 3.2: Feature point extraction with adaptation to scale change: (a) Block diagram of the feature extraction process, and (b) Illustration of the feature extraction process showing the input image, Mexican-hat wavelets, the obtained response, and the extracted points superimposed on the image. The points are represented by circles of different radii, indicating different scale associated with each point.

where

$$\Re(\mathbf{x}, s_m, s_{p_i}) = I(\mathbf{x}) \otimes Mex(\mathbf{x}, s_m, s_{p_i}) \tag{3.4}$$

represents the convolution between the image $I$ and the Mexican-hat wavelet given by Eqn. 3.2.

A computationally efficient way to implement the filtering operation in Eqn. 3.4, to obtain $\Re(\mathbf{x}, s_m, s_{p_i})$, is using the Fourier transform of the Mexican-hat wavelet given by

$$Mex_f(s_{p_i}\ \mathbf{x_f}, s_m) = s_{p_i}^2\ \sigma^3 \left((s_{p_i}x_{f_1})^2 + (s_{p_i}x_{f_2})^2\right)\ e^{\left(-\frac{1}{2}\sigma^2((s_{p_i}x_{f_1})^2+(s_{p_i}x_{f_2})^2)\right)} \tag{3.5}$$

The response $\phi(\mathbf{x}, s_1, s_2, s_{p_i})$ can be obtained in the frequency domain using:

$$\begin{aligned}
\phi(\mathbf{x}, s_1, s_2, s_{p_i}) &= |\Re(\mathbf{x}, s_1, s_{p_i}) - \Re(\mathbf{x}, s_2, s_{p_i})| \\
&= |IFFT\{I_f(\mathbf{x_f}) \times Mex_f(\mathbf{x_f}, s_1, s_{p_i}) - I_f(\mathbf{x_f}) \times Mex_f(\mathbf{x_f}, s_2, s_{p_i})\}\,| \\
&= |IFFT\{I_f(\mathbf{x_f}) \times (Mex_f(\mathbf{x_f}, s_1, s_{p_i}) - Mex_f(\mathbf{x_f}, s_2, s_{p_i}))\}\,| \qquad (3.6)
\end{aligned}$$

where $I_f(\mathbf{x_f})$ denotes the Fourier transform of $I(\mathbf{x})$ and, FFT and IFFT represent the Fourier transform and its inverse, respectively.

2. The second stage of the feature extraction process localizes the feature points of the image by finding the local maxima of the response $\phi(\mathbf{x}, s_1, s_2, s_{p_i})$. A local maximum is a point with maximum value that is greater than a specified threshold $T_n$ in a disk-shaped neighborhood of radius $s_{p_i} \times r_n$. Depending on the image content, there might be a large number of local maxima that can be detected in the feature point response. By tuning the neighborhoods parameters (block sizes, radius of neighborhoods, etc.), the closeness of the extracted feature points to each other can be controlled to ensure their stabilities and to reduce the computations required for large number of points. Further, local maxima

with low values are more sensitive to image distortions. Therefore, only local maxima that are greater than the specified threshold $T_n$ are chosen. It was found, through experiments, that a threshold $T_n$ with the value of 10% of the global maximum value gives good results.

In order to avoid the effects of the image borders on the feature point extraction process, only the image area that is away from the four image borders by a distance $s_{p_i} \times Br$ is considered for finding the local maxima . This local maxima obtaining stage will be applied to the response $\phi(\mathbf{x}, s_1, s_2, s_{p_i})$, in Eqn 3.3 using the following algorithm:

(a) Find the maximal values that are greater than the specified threshold $T_n$ in equally non-overlapped blocks of size $(s_{p_i} \times l_n) \times (s_{p_i} \times l_n)$; such initial maximal values may include points on the boundaries of the blocks, which do not represent a local maximum of $\phi(\mathbf{x}, s_1, s_2)$.

(b) Take each maximal point as the centre of a disk-shaped neighborhood of radius $s_{p_i} \times r_n$ and find one local maximum in each neighborhood; this will eliminate maximal points that are not local maxima of $\phi(\mathbf{x}, s_1, s_2)$ or local maxima that are too close to each other.

(c) Repeat step-b until the obtained local maxima do not change locations.

(d) In order to avoid the effects of the image borders on the feature extraction process, only the maxima found in the image area that is away from the image border by a distance $s_{p_i} \times Br$ are kept.

For simplicity, the radius of the neighborhoods $r_n$ is set to $0.75\ l_n$ and the border distance $s_{p_i} \times Br$ is set to $s_{p_i} \times 2\ l_n$. Thus, only one parameter is used to control the neighborhood size, that is $l_n$.

The locations of the obtained local maxima ($P_i, i = 1, 2, \ldots, K$ are taken as the extracted feature points, where $P_i = (x_{1i}, x_{2i}, s_{p_i})$ are the coordinates of a point $P_i$, $s_{p_i}$ is the associated scaling factor and $K$ is the numbers of feature points of $I$. The process of feature point extraction from an image using the technique proposed here is illustrated in Fig. 3.2 and 3.3.

In order to illustrate the details of using range of scales, Fig. 3.4 shows an input image and, for different values of the adaptation scale ($s_{p_i}$), the obtained responses along with the corresponding extracted feature points superimposed on the image.

The following remarks can be made regarding the implementation and parameter values used in the proposed method:

*Remark 1*: The Mexican-hat wavelet has perfect circular symmetry in the frequency and spatial domains as can be seen in Fig. 3.5. Its Fourier transform indicates that the Mexican-hat wavelet is equivalent to a band-pass filter [58]. Tuning the wavelet scale $s_m$ controls the spread of the wavelet in the spatial domain and the bandwidth in the frequency domain: setting $s_m$ to a low value moves the passband of the band-pass filter to lower frequency while using a high value moves the passband of the band-pass filter to higher frequency.

In order to find the best scales ($s_1$ and $s_2$), a set of 25 different images with various characteristics (camera taken photography, aerial and satellite images, etc.) are tested to investigate the effect of choosing such wavelet scales. Each image of the set is further distorted using a combination of rotation, noise contamination and brightness change. Each image, and its distorted version, are considered in the feature extraction process using several combinations of $s_1$ and $s_2$. In Fig. 3.6, the feature points repeatability rate for different values of $s_1$ and $s_2$ is shown. It can be seen in Fig. 3.6 that $(s_1 = 3, s_2 = 9)$ and $(s_1 = 2, s_2 = 4)$ give the highest repeatability rates. Further, during these experiments, the total number of feature points extracted

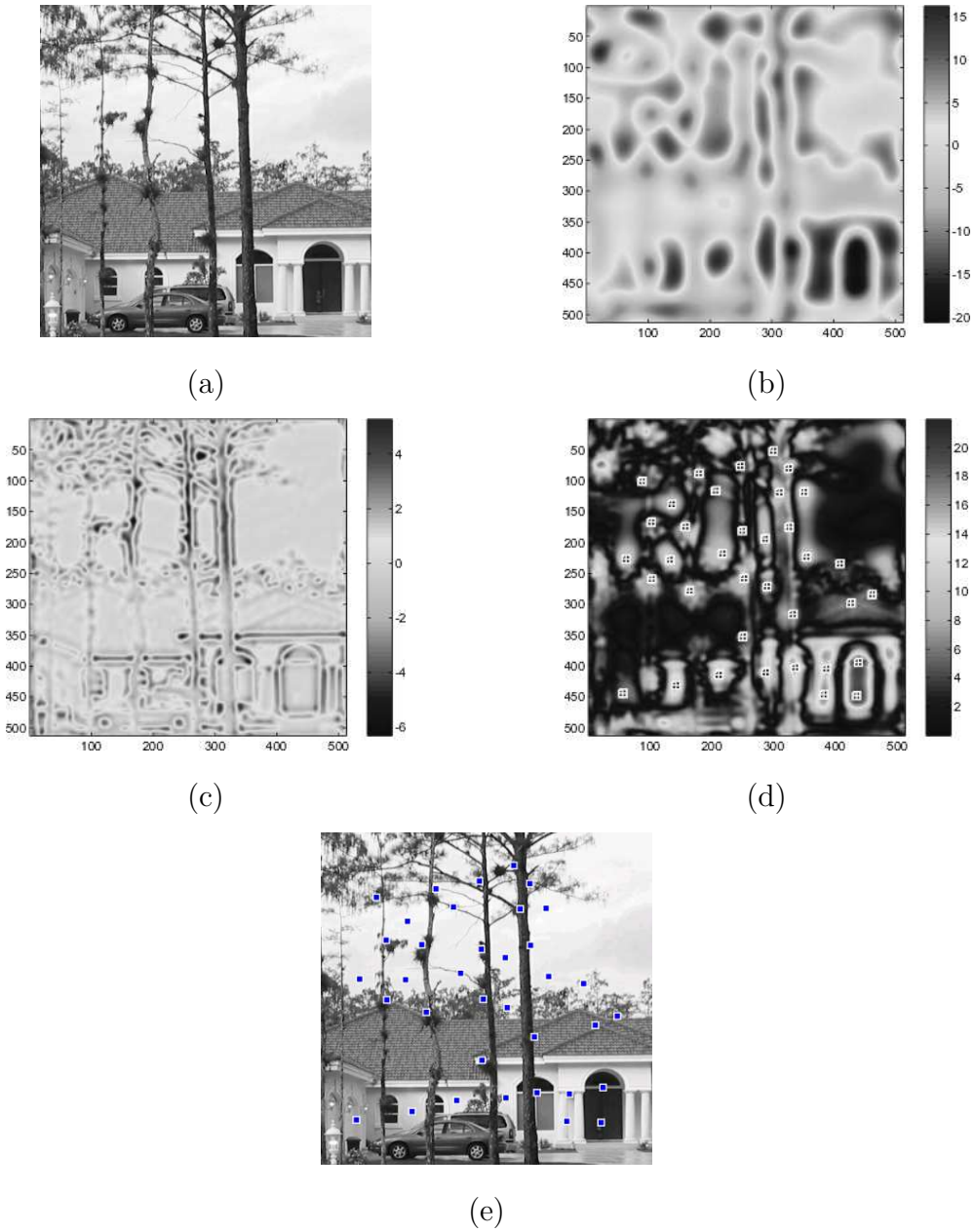



Figure 3.3: Feature point extraction stages: (a) Input image, (b) Response of applying Mexican-hat wavelet with scale 2, (c) Response of applying Mexican-hat wavelet with scale 4, (d) Absolute difference of the two responses, and (e) Input image with the extracted points superimposed on the image.

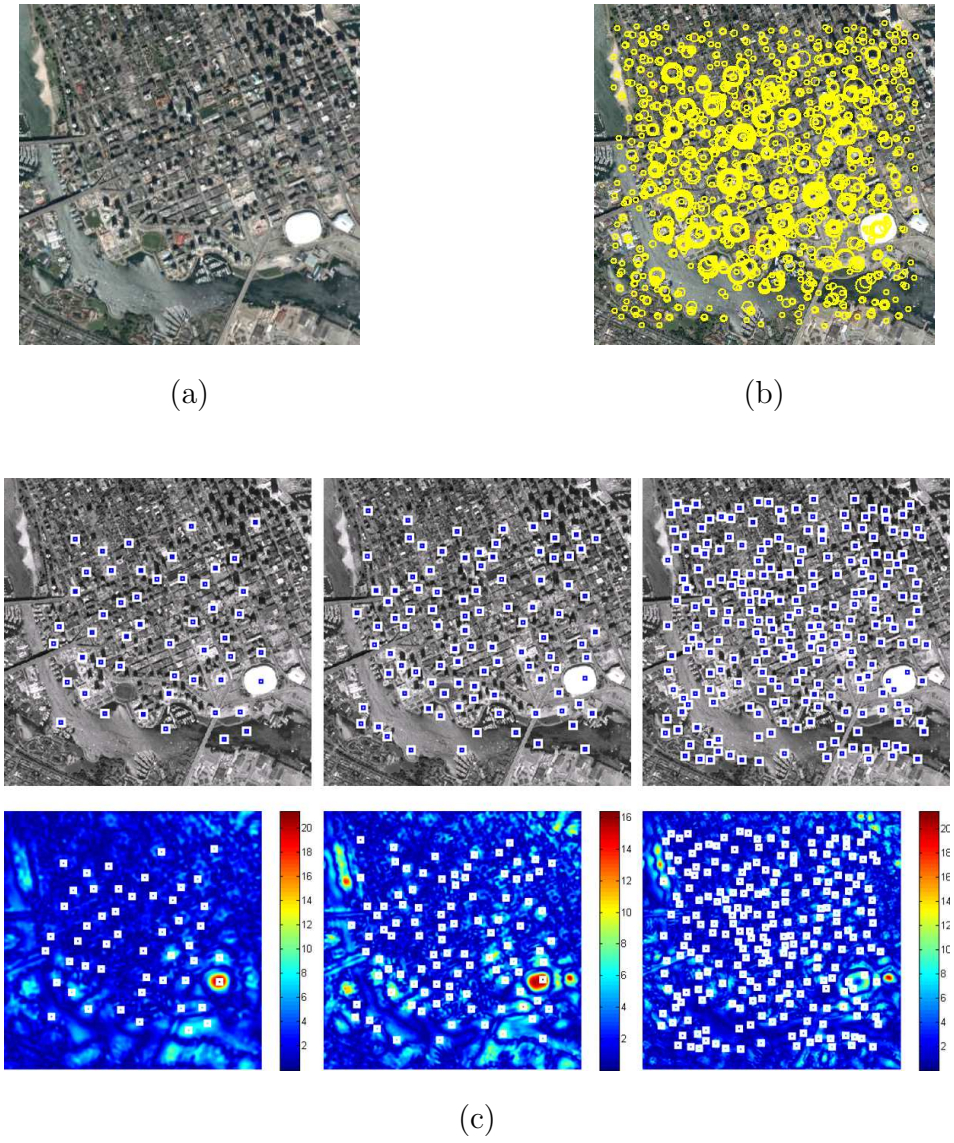


Figure 3.4: Example of feature point extraction with adaptation to scale change: (a) input image, (b) the extracted points superimposed on the image, and (c) the obtained responses along with the corresponding extracted feature points superimposed on the image, for different values of the adaptation scale ($s_{p_i}$).

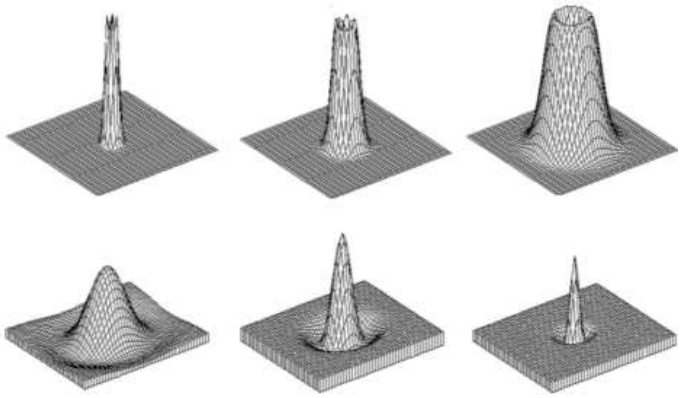


Figure 3.5: Mexican-hat wavelets for different scale parameters $s$: in the first row are the magnitude responses in the frequency domain and in the second row are the corresponding impulse responses in the spatial domain.

using $(s_1 = 2, s_2 = 4)$ was lower than the one using $(s_1 = 3, s_2 = 9)$. Since the computation time required in finding the correspondence between feature points in the two images depends on the total number of points extracted, $(s_1 = 2, s_2 = 4)$ is chosen.

*Remark 2*: To illustrate the motivation for introducing the scaling factor $s_{p_i}$ in the Mexican-hat wavelet, the effect of using $s_{p_i}$ on the feature point extraction is considered in an example given in Fig. 3.7. In the first row the original image and two scaled version of this image using scaling factors of 1.5 and 2 are presented. The feature points extraction in these images is carried out using the Mexican-hat wavelet with scaling factors $s_{p_i} = 1$, $s_{p_i} = 1.5$ and $s_{p_i} = 2$ respectively. In the second row, the same scaled images are presented as in the first row. The difference is that in the second row the feature points extraction for all three images is done using the Mexican-hat wavelet with scaling factor $s_{p_i} = 1$. The following observations can be made by comparing the locations of the extracted feature points among images in a row. In the top row, feature points are detected in the same relative locations in

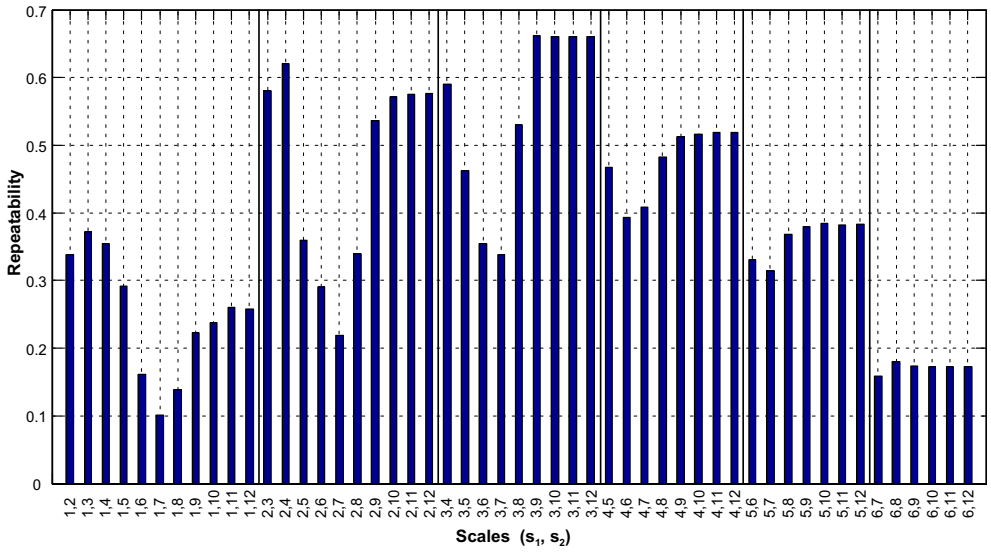



Figure 3.6: Feature point repeatability rate versus the Mexican-hat wavelet scales ($s_1$ and $s_2$).

the image in all three cases. In the second row, the feature points extracted are not corresponding to the same relative locations in the image. This indicates that the feature points extracted remain constant when the relative size of image features and width of the Mexican-hat remain unchanged. If they are changed, the feature points do not longer correspond to the same image locations. This illustrates qualitatively the effect of using a scaling factor on the Mexican-hat wavelet. This property of the feature extractor is important for image registration, presented in the next chapter.

*Remark 3*: The parameter that controls the neighborhood size, that is $l_n$, depends primarily on the size of the image and the spread of the wavelet in the spatial domain. This parameter is chosen to set the radius of the disk-shaped neighborhood $r_n$ as the radius/width of the central positive peak of the Mexican-hat wavelet with the broader spread (using $s_1 = 2$), that is 30. Hence $l_n$ used in with the value of 40 to maintain $r_n = 0.75\ l_n = 30$.

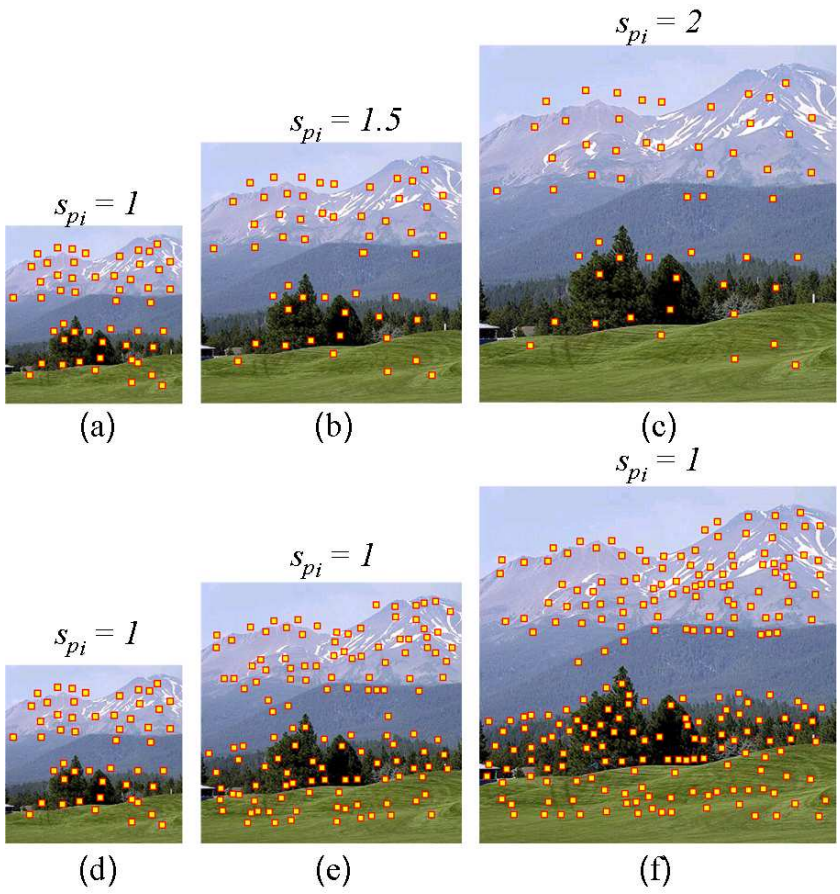



Figure 3.7: Feature point extraction in images with scale change: Images in both rows are the original image and two images scaled using scaling factors 1.5 and 2 respectively. The feature points on the top row have been extracted using the Mexican-hat wavelet with scaling factors $s_{p_i} = 1$, $s_{p_i} = 1.5$ and $s_{p_i} = 2$ respectively. The feature points in the bottom row have been extracted using the Mexican-hat wavelet with scaling factor $s_{p_i} = 1$ in all three cases.

### 3.2.3 Performance Analysis

The performance of the presented feature point extractor was evaluated using various images of different characteristics. The criteria used in evaluating the performance of the feature points extractor are based on visual inspection (subjective test), repeatability rate and localization accuracy (objective test).

- *Examples using different image distortions*

  Various images of different characteristics are used here to extract the feature points after applying several types of distortions. The criterion used here is based on visual inspection of the locations of the extracted feature points. A robust feature point extractor should results in as many as possible feature points that are in the same, or within the neighborhood of, locations, with respect to the contents of the image, regardless of any degradations or distortions applied on the image.

  An example showing the extracted feature points from different distorted versions of an image is shown in Fig. 3.8. In each case, the upper sub-figure shows the feature points superimposed on the image while the lower sub-figure shows the response of the image to the feature detection operation and the local maxima in the response are marked. It can be seen in Fig. 3.8 that many feature points detected in the original image, Fig. 3.8-a, are also detected in the same relative locations in the distorted images. Further, as an objective criterion, table 3.1 shows the number of points ($Npoint_T$, $Npoint_C$, $Npoint_O$, $Npoint_M$), which are extracted from the overlapping area of the undistorted and distorted images. $Npoint_T$ is the number of points detected in original undistorted image, $Npoint_C$ is the number of points detected in both the original and the distorted images, $Npoint_O$ is the number of points detected in the original image but not in the distorted image, and $Npoint_M$ is the number of points detected in

the distorted image but not in the original image. This indicates that large

Table 3.1: Feature point extraction of distorted/degraded images.

| Distortion/Degradation | $Npoint_T$ | $Npoint_C$ | $Npoint_O$ | $Npoint_M$ |
|---|---|---|---|---|
| (b) Blurring ($3 \times 3$ averaging filter) | 36 | 32 | 4 | 5 |
| (c) Brightness change by $-80$ (in 255-level) | 36 | 33 | 3 | 1 |
| (d) Rotation by $30^o$ | 34 | 30 | 4 | 2 |
| (e) Gaussian noise of zero-mean and standard deviation of 57 | 36 | 32 | 4 | 6 |
| (f) 'Salt & pepper' noise of density 0.1, (affects approximately 10% of the image area), and rotation by $30^o$ | 34 | 27 | 7 | 5 |

percentage of feature points are extracted in the same relative location in both the original and distorted images.

- *Comparison with other feature point extractors*

  The criterion used here is based on quantifying the repeatability rate and localization accuracy of the extracted feature points. Through a comparison with another feature point extractor, the performance of the presented feature point extractor is discussed, with regard to those criteria.

  There are many feature point extractors, which have been presented in the literature, and an interesting question is how does the performance of the Mexican-hat wavelet-based feature point extractor compares with them. One of the commonly used extractors is the Harris detector [11,59] and many techniques have been compared to it. The Harris detector will, thus, be used here for comparison. The performance evaluation will be done in terms of repeatability rate and localization accuracy.

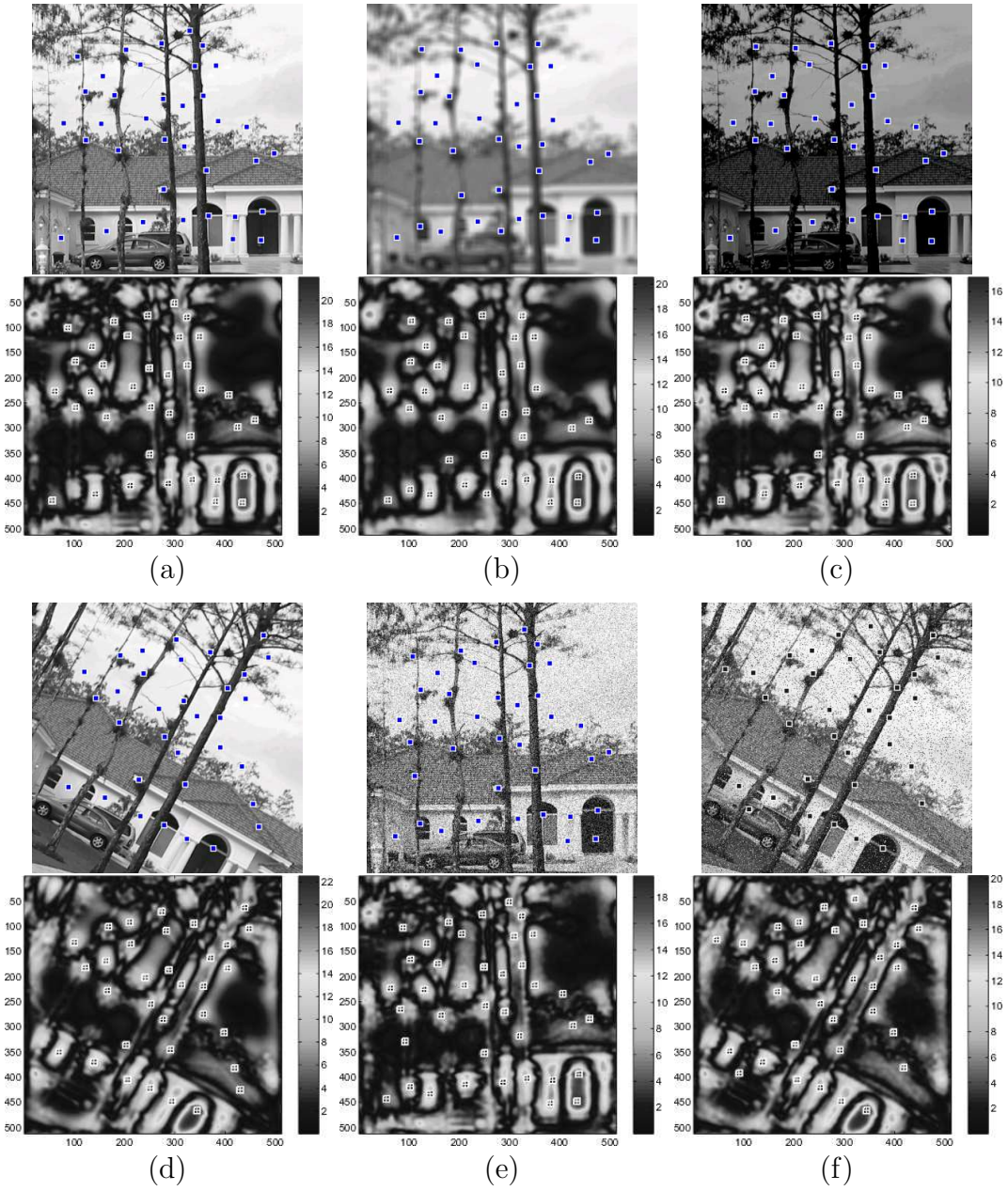



Figure 3.8: Feature point extraction of distorted/degraded images: (a) No distortion, (b) Blurring, (c) Brightness change, (d) Rotation, (e) Gaussian noise contamination, and (f) 'Salt and Pepper' noise contamination.

An image and its distorted version are considered and the feature points in each image are extracted using the Harris detector and the Mexican-hat wavelet detector. The repeatability rate is the percentage of feature points detected in the original and distorted images. A feature point is considered detected if it is detected in both images and the corresponding locations do not differ by more than two pixels.

$$Repeatability = \frac{Npoint_C}{Npoint_O + Npoint_C + Npoint_M}, \tag{3.7}$$

The localization accuracy is quantified by obtaining the mean of distance between corresponding feature points, in pixels.

The standard image 'Lena' is used as an example and the performance of each detector is tested with respect to different types of distortion. The degradations considered are JPEG compression, Gaussian noise addition and impulsive noise ('Salt and Pepper' noise) addition. Comparison results that illustrate the repeatability rate and localization accuracy of the two feature point detectors are shown in Fig. 3.9. Other images used for comparison lead to similar results. The comparison results indicate that the Mexican-hat wavelet-based feature point detector has shown better repeatability rate for all tested images. Further, both of the detectors have comparable localization accuracy for most images and in some images, the Mexican-hat wavelet-based feature point detector has better localization accuracy. Based in this comparison, the Mexican-hat wavelet-based feature point detector can perform as well or better than the Harris detector.

The scale information can be used in the feature point extraction process to accommodate the scale change of images and can be utilized in applications dealing with images of different scales, as can be seen in our proposed image registration technique in the subsequent chapter. The performance of the

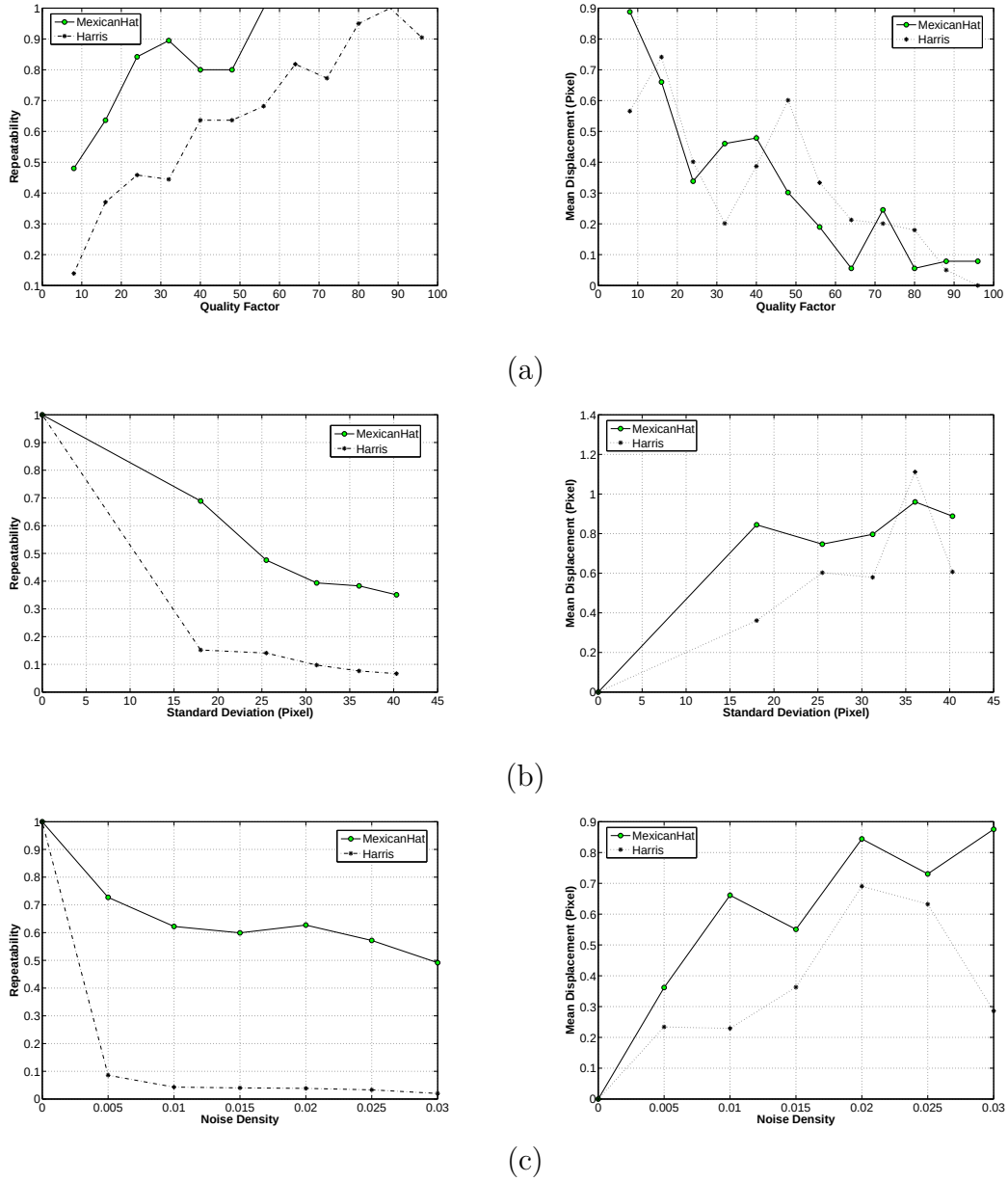



Figure 3.9: Feature point repeatability (left) and localization accuracy (right) for several types of distortions for image 'Lena': (a) JPEG compression, (b) Gaussian noise and (c) 'Salt and Pepper' noise.

presented feature point extractor, taking into consideration the different scales issue, is compared with other feature point extractors with regard to the overall performance of the used application, image registration in our case, and discussed in next chapter.

## 3.3 Conclusions

In this chapter, a feature point extractor that is based on scale-interaction of Mexican-hat wavelets with the ability to handle images with scale change was introduced. In order to adapt the feature point extraction process to images with scale change, a range of scaling factors is used for the Mexican-hat wavelet and feature points for different scaling factors are extracted. This property of the feature point extractor is important for image registration, presented in Chapter 4.

The presented feature points extractor was tested using a various images of different characteristics and its performance was evaluated using visual inspection (subjective test), repeatability rate and localization accuracy (objective test). Experimental results indicate that:

- The locations of feature points extracted using the proposed technique have reasonable invariance to image rotation, translation and scale change.
- The extracted feature points are robust against lossy image compression, image blurring and noise contamination.

The presented feature point extraction method will be used in image registration and normalization methods, as discussed in the subsequent chapters.

# Chapter 4

# A Robust Image Registration Technique using Feature Point Extraction and Zernike Moment Invariants

This chapter introduces an algorithm for image registration. The main objective of the proposed algorithm is accurately registering images with partial overlap that are geometrically distorted and in addition they may have undergone degradations caused by, for example, noise contamination, blurring, etc. The geometric distortions considered in the registration process are the global $2D$ affine transformations. The proposed algorithm consists of three main steps: extracting feature points using a feature point extractor based on scale-interaction of Mexican-hat wavelets, obtaining the correspondence between the feature points of the reference and the target images based on Zernike moments of neighborhoods centered on the feature points, and estimating the transformation parameters between the two images using an iterative weighted least squares algorithm.

This chapter is organized as follows. Section 4.1 discusses the proposed registration algorithm: its objectives and requirements. Section 4.2 describes the

proposed registration algorithm in detail. In Section 4.3, experimental results are presented and the performance of the proposed algorithm is discussed. Finally, Section 4.4 summarizes the chapter.

## 4.1 Introduction

Image registration has found applications in numerous real-life applications such as remote sensing, medical image analysis, computer vision and pattern recognition [3]. Given two, or more, images to be registered, image registration estimates the parameters of the geometric transformation model that maps a given target image to the reference one.

### 4.1.1 Objectives

This chapter introduces an algorithm for image registration that aims at accurately registering images that are geometrically distorted and, in addition, they may have undergone additional degradations. The proposed algorithm is required to satisfy the following objectives:

- The principal objective is to develop an image registration method for images of different characteristics, for example, aerial images, satellite images and digital camera photography.
- It should be capable of dealing with significant geometric distortions that involve global $2D$ affine transformations, and when there is only partial overlap between the images.
- It should have a reasonable robustness to possible degradations such as noise contamination, blurring, etc.

- The registration algorithm should be automatic, which means that the procedure should not require any manual steps.

### 4.1.2 Geometric Distortions Considered in The Registration Process

The types of distortions considered here are the global $2D$ affine transformations. A combined transformation of these types typically has six parameters: translation parameters in $x$ and $y$ directions ($t_x$ and $t_y$ respectively) and four affine transformations parameter ($a_{11}, a_{12}, a_{21}$ and $a_{22}$). This transformation maps a point $\mathbf{x} = (x_1, x_2)$ of an image $I$ to a point $\mathbf{x}^{'} = (x_1^{'}, x_2^{'})$ of the transformed image $I^{'}$ as follows:

$$\begin{aligned} \mathbf{x}^{'} = f(\mathbf{x}, \mathbf{z}) = \begin{pmatrix} x_1^{'} \\ x_2^{'} \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} t_{x1} \\ t_{x2} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + t_{x1} \\ a_{21}x_1 + a_{22}x_2 + t_{x2} \end{pmatrix}, \end{aligned} \tag{4.1}$$

where $f$ is the transformation function, $\mathbf{z} = [a_{11} \quad a_{12} \quad a_{21} \quad a_{22} \quad t_{x1} \quad t_{x2}]^T$ is a vector of the transformation parameters and $I(\mathbf{x})$ represents a pixel value at location $\mathbf{x} = (x_1, x_2)$.

## 4.2 The Proposed Registration Algorithm

The problem of image registration is to estimate the transformation parameters, as in Eqn. 4.1 using the reference and the target images. The image quality may be, further, degraded, for example by noise contamination, or the two images may have only partial overlap between them.

The proposed algorithm involves three stages: extracting of feature points, obtaining the correspondence between the feature points of the two images, and estimating the transformation parameters. Typically, the feature extraction process

and obtaining the correspondence between feature points are carried out on gray-scale images or on the luminance component of color images.

The feature point extractor, presented in Chapter 3, is utilized here to extract two sets of feature points from the reference and the target images respectively. These feature points are scale dependent and the proposed technique includes adjusting the Mexican-hat wavelet size to obtain different feature points for different scales.

The correspondence between the feature points from the two images is evaluated using Zernike moment invariants of circular neighborhoods centered on the feature points. Zernike moments have been evaluated and shown to give very good results compared to other types of moments with respect to information redundancy and sensitivity to noise [43]. Another important property of Zernike moments is that their magnitude is rotationally invariant [60].

The transformation parameters are estimated using an adaptive weighted least squares technique with an objective function that depends on the weighted difference between the locations of a set of feature points in the reference image and another set in the target image. Using this approach, outliers are eliminated and inliers are weighted according to the relative magnitude of their residuals.

The steps of the proposed registration algorithm are shown schematically in Fig. 4.1 and described in detail in the following subsections.

### 4.2.1 Feature Point Extraction

In order to extract two sets of feature points from the given two images to be registered, the feature point extractor, presented in Chapter 3, is used. This feature point extraction is applied to both images ($I$ and $I^{'}$) and the locations of the obtained feature points are found as ($P_i, i = 1, 2, \ldots, K$ for $I$ and $P_i^{'}, i = 1, 2, \ldots, K^{'}$ for $I^{'}$), where $P_i = (x_{1i}, x_{2i}, s_{p_i})$ are the coordinates of a point $P_i$, $s_{p_i}$ is the associated scaling

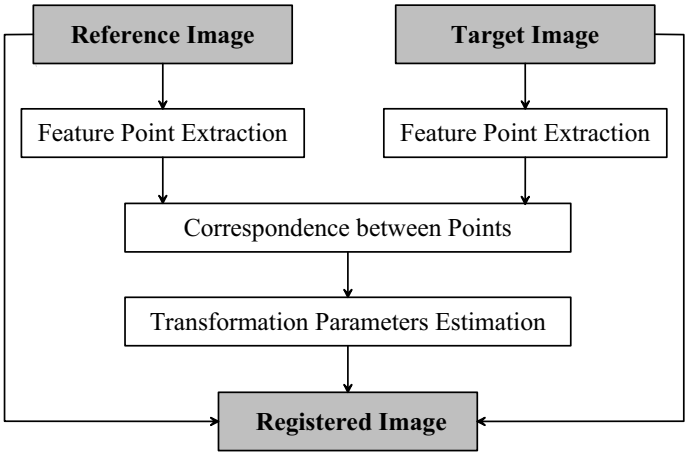



Figure 4.1: Block diagram of the proposed registration algorithm.

factor and, $K$ and $K^{'}$ are the numbers of feature points of $I$ and $I^{'}$, respectively. An example illustrating the extracted feature points from two different images is shown in Fig. 4.2.

### 4.2.2 Correspondence between Points

The feature extraction process discussed in the previous sub-section applied on the two images results in two sets of feature points, $P$ and $P^{'}$. The number of the feature points in the reference image is $K$ and the corresponding number of the target image is $K^{'}$. The objective of this step is to pair feature points of the reference image with the corresponding ones of the target image. This is done using similarity measures based on Zernike moments-based descriptors [43, 60, 61] using circular neighborhoods centered on each feature point $P_i$ in the reference image and each point $P_i^{'}$ in the target image. The reason for using circular neighborhoods is that the magnitudes of the complex Zernike moments are rotational invariant. Further, the Zernike moments of an image $I$ are related to the Zernike moments of the resized image $I^{'}$, obtained

(a) (b)

Figure 4.2: Feature point extraction from two images: the extracted points superimposed on the two images in (a) and (b). The points are represented by circles of different radii, indicating different scale associated with each point.

by scaling $I$ using a scaling factor $s_{p_i}$, using the following equation [62]

$$Z_{p,q} = \frac{1}{s_{p_i}^2} Z'_{p,q} \tag{4.2}$$

The correspondence between feature points in the two images is obtained using the following algorithm:

1. For each point of $P$ and $P^{'}$ in images $I$ and $I^{'}$, respectively, take a circular neighborhood of radius $s_{p_i} \times r_v$ and construct a descriptor vector $P_d$ as

$$P_d = \frac{1}{s_{p_i}^2}[\,|Z_{1,1}|, \ldots, |Z_{p,q}|, \ldots, |Z_{10,10}|\,], \tag{4.3}$$

where $|Z_{p,q}|$ is the magnitude of Zernike moment of a non-negative integer order $p$, $p-|q|$ is even, and $|q| \leq p$. When computing the Zernike moments of a circular neighborhood located around a feature point, the feature point is taken as the origin and the coordinates of each pixel inside the neighborhood are mapped to

the range inside a unit circle, i.e., $\sqrt{\hat{x_1}^2 + \hat{x_2}^2} \leq 1$. Zernike moments of order $p$ are given by

$$Z_{pq} = \frac{(p+1)}{\pi} \sum_{\hat{x_1}} \sum_{\hat{x_2}} V_{pq}^*(r, \theta) A(\hat{x_1}, \hat{x_2}), \tag{4.4}$$

where

$$r = \sqrt{\hat{x_1}^2 + \hat{x_2}^2}, \quad \theta = \tan^{-1}(\hat{x_2}/\hat{x_1}), \tag{4.5}$$

and $A(\hat{x_1}, \hat{x_2})$ represents the intensity at a pixel inside the circular neighborhood. In the above equation, $V_{pq}^*$ denotes the complex conjugate of the Zernike polynomial of order $p$ and repetition $q$ which can be defined as

$$V_{pq}(r, \theta) = R_{pq}(r) e^{jq\theta}, \tag{4.6}$$

where $R_{pq}(r)$ is a real-valued radial polynomial defined as

$$R_{pq}(r) = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s \quad (p-s)! \quad r^{p-2s}}{s! \left(\frac{p-2s+|q|}{2}\right)! \left(\frac{p-2s-|q|}{2}\right)!}, \tag{4.7}$$

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

While higher order moments contain information about fine details in the image, they are more sensitive to noise than lower order moments [43]. Therefore, the highest moment order used in the descriptor vector $P_d$ (10 in the algorithm) is chosen to achieve a compromise between noise sensitivity and the information content of the moments. If the image (or region of interest, i.e., the circular neighborhoods) is rotated by an angle $\varphi$, then the Zernike moment $Z_{pq}$ of the rotated image can be obtained as

$$Z'_{pq} = Z_{pq} e^{-jq\varphi}, \tag{4.8}$$

Thus, the magnitudes of the Zernike moments can be used as rotationally invariant image features.

In order to illustrate such Zernike moment-based descriptors, several distorted versions of an image are used to compute the descriptor vectors for the same neighborhood in each, denoted by the shown circles in Fig. 4.3-a-(ii–vi). These descriptor vectors are plotted in Fig. 4.3-b. It can be see that the computed descriptor are similar regardless the image degradation due to the applied distortions and this demonstrate their less sensitivity to noise and distortions, which make them effective in obtaining the correspondence between points. On the other hand, different neighborhood is used to compute the descriptor vectors, as shown in Fig. 4.3-a-vii, and it can be seen in Fig. 4.3-b that the computed descriptors are distinguishable for different neighborhoods.

Another example showing the computed descriptors of two corresponding neighborhoods in images with scale change, by considering the scale adjustment when computing descriptors, is shown in Fig. 4.4. Clearly, the effect of that adjustment can be noticed in Fig. 4.4-b from the similarity of the computed descriptors.

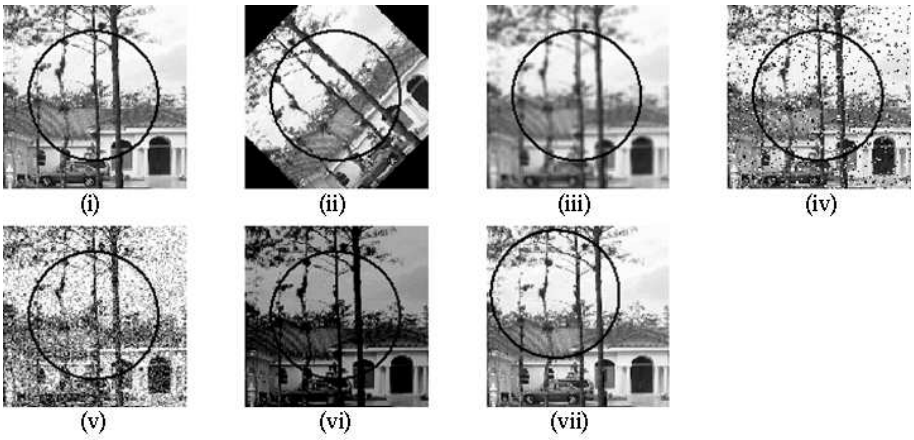
2. Construct the distance matrix $C$, where each entry $c_{ij}$ of this matrix is given by

$$c_{ij} = \ell_1(P_{d_i} - P'_{d_j}) = \sum_{m=1}^{n} |P_{d_i}(m) - P'_{d_j}(m)|, \tag{4.9}$$

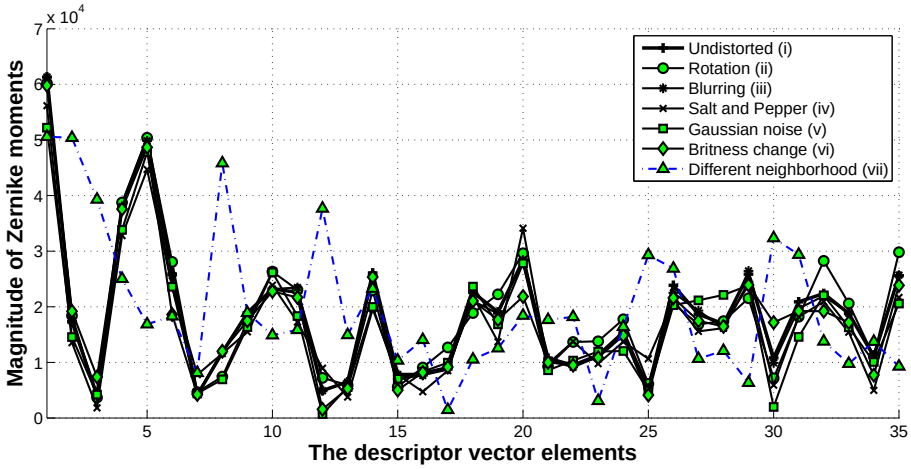
where $P_{d_i}(m)$ and $P'_{d_j}(m)$ are the entries of $P_{d_i}$ and $P'_{d_j}$, respectively, $i = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, K'$. In other words, each entry $c_{ij}$ represents the $\ell_1 - norm$ of the difference between the two descriptor vectors of the feature points $P_i$ and $P'_j$ in the reference and the target images, respectively.

3. In the distance matrix $C$, find the minimum distance coefficients along rows and along columns and denote the indices of the corresponding entries of $C$ as:

$$row_i \;=\; index\{\min_j\{c_{ij}\}\},\ and$$

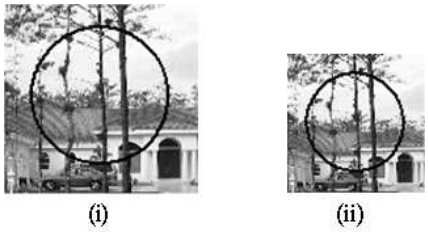



(a)

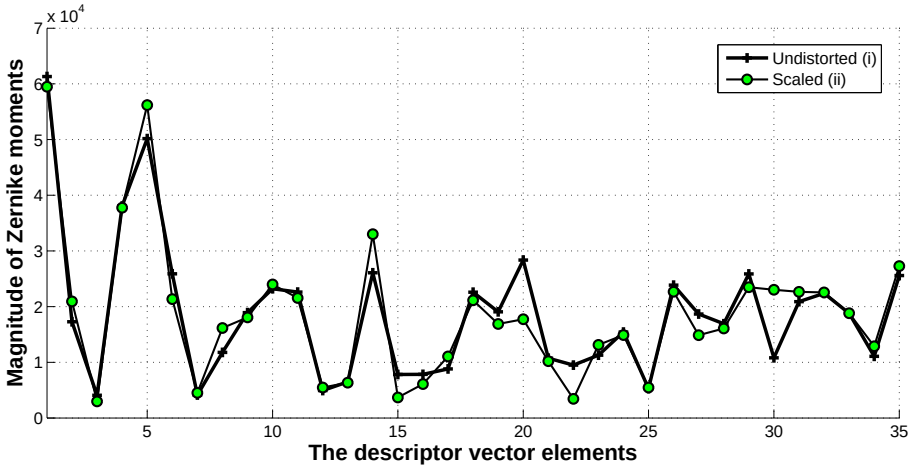



(b)

Figure 4.3: The computed descriptors for a neighborhood in an image, after applying several distortions on that image: (a) shows the distorted versions of the image and (b) shows a plot of the computed descriptors.




(a)




(b)

Figure 4.4: The computed descriptors for two neighborhoods in images with scale change: (a) shows the images and (b) shows a plot of the computed descriptors.

$$col_j \quad = \quad index\{\min_i\{c_{ij}\}\}. \tag{4.10}$$

4. A correspondence between two points $P_i$ and $P_j^{'}$ is established if, and only if, $row_i = col_j$, i.e., only if the minimum distance coefficient in a row is also the minimum distance coefficient in the associated column of $C$. This results in $\hat{K}$ paired points, where $\hat{K} \leq \min(K, K^{'})$.

*Remark 1*: When registering images that involve scaling, the used feature point extractor relies on finding feature points in both images at different scaling factors and match the feature points and their appropriate scaling factors. Since it is not practical to extract feature points at all scales, scale quantization is necessary in the scale range $s_{p_i}$. Using a large value for $s_{p_i}$ enlarges the Mexican-hat wavelets size and, consequently, fewer points are detected, as the response $\phi$ gets smoother. Therefore, $s_{p_{max}}$ is obtained as the largest value of $s_{p_i}$ for which feature points can be detected. Using a small value for $s_{p_i}$ makes the Mexican-hat wavelets width small and its response moves toward detecting image edges. After several experiments $s_{p_{min}}$ with the value of 0.3 has been chosen. In order to test the impact of the scale range step size $s_{p_{step}}$ several experiments were conducted using a variety of images. The number of correct feature point correspondences for different step sizes and the number of filtering operations for different step sizes are shown in Fig. 4.5. It can be seen from that figure that smaller step sizes lead to more correct correspondences while at the same time the number of filtering levels is increased. More correct correspondences between points helps in estimating the transformation parameters at the cost of more computations for finding such correspondences between points. A good compromise can be obtained by using a scale range step $s_{p_{step}} = 0.1$, which maintains a reasonable number of correct correspondences between points.

*Remark 2*: The parameter $r_v$, which used in computing the Zernike moments, should

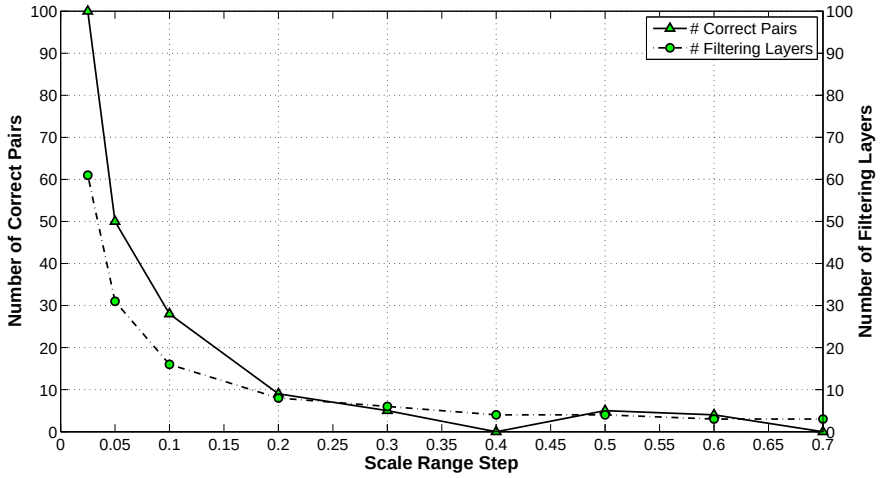


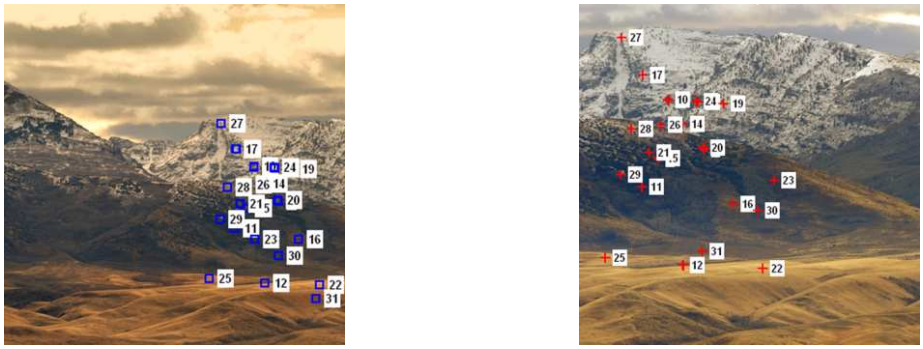

Figure 4.5: Number of Correct Feature Point Pairs versus $s_{p_{step}}$ and Number of Filtering operations versus $s_{p_{step}}$.

be chosen such that the neighborhood of each feature point contains enough details to construct an effective descriptor vector. Clearly the larger the neighborhood's size, the more image details it contains and consequently the better correspondence can be obtained but larger neighborhoods require more computations. A good compromise value has been found to be 45 pixels for $r_v$.

An example illustrating the obtained correspondences between the feature points of the images in Fig. 4.2 is shown in Fig. 4.6.

### 4.2.3 Transformation Parameters Estimation

The transformation parameters, required to transform the distorted image to its proper form, will be estimated by solving an iterative weighted least squares minimization problem where the objective function depends on the distance between the feature point pairs in the two images. The objective function is defined in terms



(a) (b)

Figure 4.6: Correspondences between the feature points. The paired feature points are superimposed on each image. Each point and its corresponding one in the other image are labeled with the same number.

of $\ell_2 - norm$ of the weighted errors

$$\Psi(\mathbf{z}) = \ell_2^2(W[f(P^{'}, \mathbf{z}) - P]) = \sum_{i=1}^{\hat{K}} w_i \left\| f(P_i^{'}, \mathbf{z}) - P_i \right\|^2, \tag{4.11}$$

where $W = [w_1 \dots w_i \dots w_{\hat{K}}]$, $w_i$ is the weight associated with the distance between the feature point pairs ($P_i$ and $P_i^{'}$) and $\mathbf{z} = [a_{11} \quad a_{12} \quad a_{21} \quad a_{22} \quad t_{x1} \quad t_{x2}]^T$ is a vector of the transformation parameters. The values of the weights will be determined iteratively, as discussed later. The transformation parameters can be obtained then by solving the optimization problem

$$\min_{\mathbf{z}} \Psi(\mathbf{z}) \tag{4.12}$$

Using Eqn. 4.1 and 4.11, the objective function in Eqn. 4.11 can be expressed as

$$\Psi(\mathbf{z}) = \sum_{i=1}^{\hat{K}} w_i[(a_{11}x_{1i}^{'} + a_{12}x_{2i}^{'} + t_{x1} - x_{1i})^2 + (a_{21}x_{1i}^{'} + a_{22}x_{2i}^{'} + t_{x2} - x_{2i})^2], \tag{4.13}$$

which is a quadratic form in six variables.

The gradient $g(\mathbf{z})$ of the objective function $\Psi(\mathbf{z})$ can be found as

$$g(\mathbf{z}) = \nabla\Psi(\mathbf{z}) = [\frac{\partial\Psi(\mathbf{z})}{\partial a_{11}} \quad \frac{\partial\Psi(\mathbf{z})}{\partial a_{12}} \quad \frac{\partial\Psi(\mathbf{z})}{\partial a_{21}} \quad \frac{\partial\Psi(\mathbf{z})}{\partial a_{22}} \quad \frac{\partial\Psi(\mathbf{z})}{\partial t_{x1}} \quad \frac{\partial\Psi(\mathbf{z})}{\partial t_{x2}}]^T, \tag{4.14}$$

and this expression can be rewritten in the form of

$$g(\mathbf{z}) = A\mathbf{z} + b, \tag{4.15}$$

where

$$A = 2\begin{pmatrix} \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'2} & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} x_{2i}^{'} & 0 & 0 & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} & 0 \\ \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} x_{2i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{2i}^{'2} & 0 & 0 & \sum_{i=1}^{\hat{K}} w_i x_{2i}^{'} & 0 \\ 0 & 0 & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'2} & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} x_{2i}^{'} & 0 & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} \\ 0 & 0 & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} x_{2i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{2i}^{'2} & 0 & \sum_{i=1}^{\hat{K}} w_i x_{2i}^{'}) \\ \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{2i}^{'} & 0 & 0 & \sum_{i=1}^{\hat{K}} w_i & 0 \\ 0 & 0 & \sum_{i=1}^{\hat{K}} w_i x_{1i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{2i}^{'} & 0 & \sum_{i=1}^{\hat{K}} w_i \end{pmatrix} \tag{4.16}$$

$$b = -2\begin{pmatrix} \sum_{i=1}^{\hat{K}} w_i x_{1i} x_{1i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{1i} x_{2i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{2i} x_{1i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{2i} x_{2i}^{'} & \sum_{i=1}^{\hat{K}} w_i x_{1i} & \sum_{i=1}^{\hat{K}} w_i x_{2i} \end{pmatrix}^T \tag{4.17}$$

Hence, the solution of the least-squares problem in Eqn. 4.12 can be reduced to solving a set of linear equations by solving $g(\mathbf{z}) = 0$, leading to

$$\mathbf{z}^* = -A^{-1}b \tag{4.18}$$

The matrix $A$ is invertible and Eqn. 4.18 has a unique solution if there are at least three distinct point pairs, i. e., $\hat{K} \geq 3$.

The solution of this optimization problem would give the correct transformation parameters provided that the feature point correspondence obtained is correct for all feature point pairs. This will not be the case if, for example, the two images have only partial overlap between them, or if some of the feature point pairs are less accurate than the others. An effective way to identify and eliminate such points (outliers) is using a common approach from robust statistics [63], which involves weights $w_i$ that are updated iteratively. The weights $w_i$ are computed based on the residuals

values, such that those point pairs giving large residuals are heavily weighted while large residuals are weighted lightly, or even rejected in the next iteration. There are several methods of weight updating in the literature [63]. These methods have different properties with respect to the breakdown point (the percentage of outliers that cause a wrong estimation), complexity and the convergence rate. Usually there is a trade-off between convergence rate and breakdown point [64]. The use of the mean-based function in Eqn.(4.20) leads to fast convergence to the correct parameters in the case of small percentage of outliers which has been mostly the case using the proposed algorithm. For cases with high percentage of outliers, different weight updating methods, such as the one based on the median absolute deviation [63, 64] lead to converge for cases that have up to 40% outliers.

The transformation parameters are estimated by using an iterative weighted least squares algorithm that is described by Algorithm 1 and the obtained transformation parameters are then used to transform the image to its appropriate form.

## 4.3 Experimental Results

In this section, the performance of the proposed registration algorithm is illustrated, evaluated and compared with other techniques using real-life images such as satellite images and images taken with digital cameras. These images have various different characteristics, several types of image degradations and partial overlap.

The performance of the proposed technique is evaluated using the errors in the feature point's locations of the registered images, after transformation. Suppose the coordinates of a pair of feature point in the reference and a given image are denoted by $(x_{1i}, x_{2i})$ and $(x'_{1i}, x'_{2i})$, respectively. The distance $D_i$ between the reference feature point and its registered correspondent, after transformation (using the parameters

**Algorithm 1** Iterative Weighted Least Squares Minimization

1: Find an initial estimate of the transformation parameters vector $\mathbf{z}^{(1)}$ by using Eqn. 4.18 with weights $w_i^{(1)} = 1$, $i = 1, 2, \ldots, \hat{K}$.

2: **repeat**

3: Compute the residuals for each feature point pair using

$$\Delta_i^{(n)} = \left\| f(P_i^{'}, \mathbf{z}^{(n)}) - P_i \right\|^2, i = 1, 2, \ldots, \hat{K} \tag{4.19}$$

4: Update the weights based on the values of the residuals using

$$w_i^{(n)} = \begin{cases} 0 & \text{if } \Delta_i^{(n-1)} \geq \overline{\Delta^{(n-1)}} \\ \left[1 - (\frac{\Delta_i^{(n-1)}}{\overline{\Delta^{(n-1)}}})^2\right]^2 & \text{otherwise} \end{cases} \tag{4.20}$$

where $\overline{\Delta^{(n-1)}} = Mean(\Delta^{(n-1)})$

5: Find a new solution for $\mathbf{z}$ by using Eqn. 4.18 with the updated weights.

6: **until** $|\Psi^{(n)}(\mathbf{z}) - \Psi^{(n-1)}(\mathbf{z})| < Threshold$

or a certain number of iterations has been reached

$*$ $X^{(n)}$ denotes the value of $X$ in the current iteration while $X^{(n-1)}$ implies its value in the previous iteration.

$a_{11},\ a_{12},\ a_{21},\ a_{22},\ t_{x1}$ and $t_{x2}$), can be determined as

$$D_i = \left((a_{11}x'_{1i} + a_{12}x'_{2i} + t_{x1} - x_{1i})^2 + (a_{21}x'_{1i} + a_{22}x'_{2i} + t_{x2} - x_{2i})^2\right)^{\frac{1}{2}}, \qquad (4.21)$$

The geometric registration error can be quantified in terms of the *mean* $D_M$ and the *standard deviation* $D_{STD}$ of the errors for all point pairs, which can be determined as

$$\begin{aligned} D_M &= \frac{1}{\widetilde{K}} \sum_{i=1}^{\widetilde{K}} D_i \\ D_{STD} &= \left( \frac{1}{\widetilde{K} - 1} \sum_{i=1}^{\widetilde{K}} (D_i - D_M)^2 \right)^{\frac{1}{2}}, \end{aligned} \qquad (4.22)$$

where $\widetilde{K}$ is the total number of point pairs, after excluding the pairs of false correspondence with associated weights of approximately zero value.

### 4.3.1 Registration Examples

- **Example 1**: Fig. 4.7 shows an example of two photographs [1] of an outdoor scene taken at different times. The two photographs are registered using the proposed technique. In the first step, feature extraction results in two sets of feature points for each image. These feature points are represented as circles whose centers are the coordinates of the extracted feature points, as it can be seen in Fig. 4.7-a and b. The different diameters of the circles indicate the different adaptation scales $s_{p_i}$ for each feature point. In the second step, the correspondence between the feature points in the two images is obtained and shown in Fig. 4.7-c and d. It can be seen that many feature points have been eliminated and only points for which correspondents were found are kept. Each point and its corresponding one in the other image are labeled with the same number. It can be seen that 31 feature point pairs have been matched.

[1] Available at http://www.pbase.com/ckuhn55/wyoming

All pairs except pairs #23, #30 and #31 have correct correspondence. These false correspondences are mainly due to the self-similarity of some parts of each image. Using these 31 pairs of feature points the iterative weighted least squares algorithm is used in the third step to obtain the transformation parameters. The values of the weights and residuals are given in Table 4.1. It can be seen, in this table, that after four iterations, the algorithm converges and the outliers (pairs #23, #30 and #31) are rejected since they have been assigned zero weights. The target image is transformed using the obtained transformation parameters and overlaid on the reference image as can be seen in Fig. 4.7-e. Bilinear interpolation is used in the geometric transformation. Fig. 4.7-f shows the correspondence of the feature point pairs. The crosses represent the location of the feature points of the reference image while the squares represent the location of the feature points of the target image after they are registered (geometrically transformed) and overlaid on the feature points of the reference image. As can be seen in Fig. 4.7-f, the obtained correspondence are correct, i.e., the crosses coincide with the squares except for the three false correspondences which have been eliminated by the iterative weighted least squares algorithm. The quality of registration can be evaluated using the registration error ($D_M$ and $D_{STD}$) which is computed as 0.86 and 0.46 pixel, respectively. Further, visual inspection confirms the quality of registration by the continuity of the mountain edges at the borders of the overlaid images.

- **Example 2**: Fig. 4.8 shows another example of registering two images of an outdoor view taken using a digital camera. The focal length, the position and the orientation of the camera were different for each image. The upper row of Fig. 4.8 shows the two images. Fig. 4.8-c shows the registered image overlaid on the reference image. The correspondence between the feature points and the

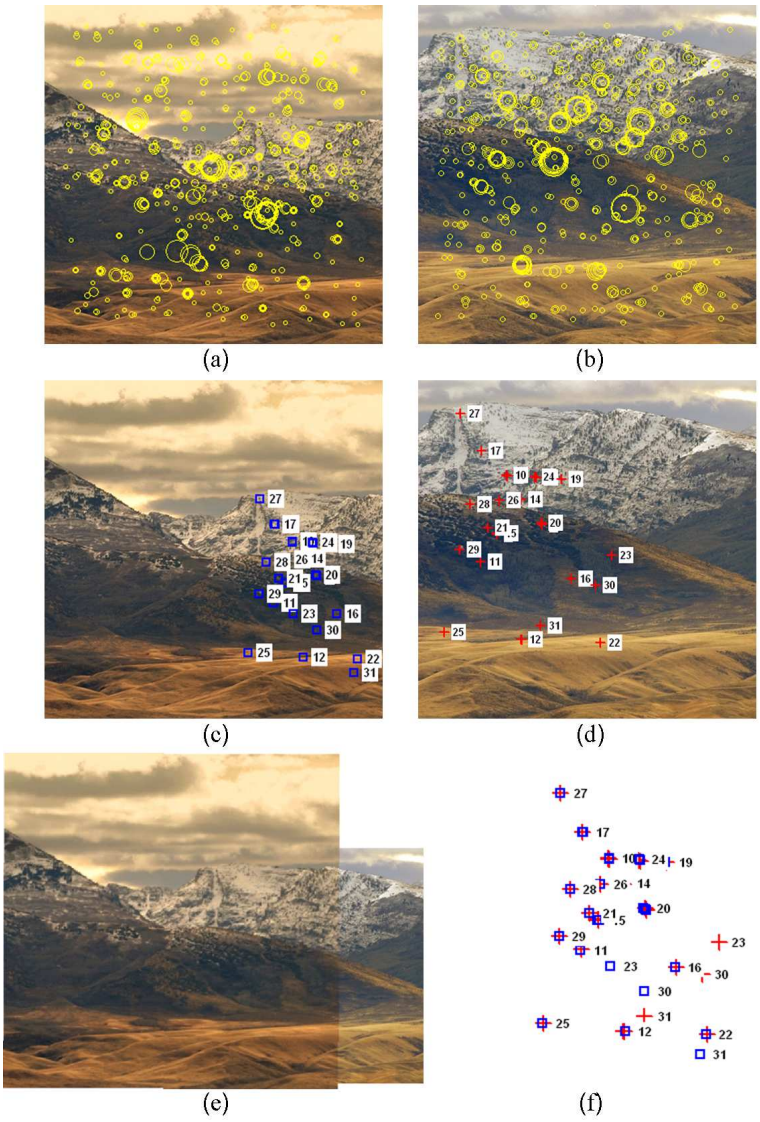



Figure 4.7: Example 1: registration of two photographs of an outdoor scene taken at different times. (a) and (b) The two images with the extracted feature points superimposed on each image; (c) and (d) The paired feature points superimposed on each image; (e) The registered image overlaid on the reference image; (f) The feature points of registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

Table 4.1: Values of residuals and weights for feature point pairs in Example 1, obtained using Algorithm 1.

| Pair No. | $w_i^{(1)}$ | $\Delta_i^{(1)}$ | $w_i^{(2)}$ | $\Delta_i^{(2)}$ | $w_i^{(3)}$ | $\Delta_i^{(3)}$ | $w_i^{(4)}$ | $\Delta_i^{(4)}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.0 | 8.10 | 0.9998 | 0.36 | 1.0 | 0.36 | 1.0 | 0.36 |
| 2 | 1.0 | 135.69 | 0.9504 | 0.34 | 1.0 | 0.33 | 1.0 | 0.33 |
| 3 | 1.0 | 14.00 | 0.9995 | 0.22 | 1.0 | 0.22 | 1.0 | 0.22 |
| 4 | 1.0 | 12.58 | 0.9996 | 0.10 | 1.0 | 0.10 | 1.0 | 0.10 |
| 5 | 1.0 | 29.25 | 0.9977 | 0.98 | 1.0 | 1.00 | 1.0 | 1.00 |
| 6 | 1.0 | 135.69 | 0.9504 | 0.34 | 1.0 | 0.33 | 1.0 | 0.33 |
| 7 | 1.0 | 14.00 | 0.9995 | 0.22 | 1.0 | 0.22 | 1.0 | 0.22 |
| 8 | 1.0 | 9.96 | 0.9997 | 1.36 | 1.0 | 1.37 | 1.0 | 1.37 |
| 9 | 1.0 | 150.46 | 0.9392 | 2.26 | 1.0 | 2.22 | 1.0 | 2.22 |
| 10 | 1.0 | 12.95 | 0.9995 | 0.24 | 1.0 | 0.25 | 1.0 | 0.25 |
| 11 | 1.0 | 47.82 | 0.9938 | 1.21 | 1.0 | 1.21 | 1.0 | 1.21 |
| 12 | 1.0 | 150.46 | 0.9392 | 2.26 | 1.0 | 2.22 | 1.0 | 2.22 |
| 13 | 1.0 | 42.01 | 0.9952 | 2.90 | 1.0 | 2.91 | 1.0 | 2.91 |
| 14 | 1.0 | 5.35 | 0.9999 | 0.46 | 1.0 | 0.46 | 1.0 | 0.46 |
| 15 | 1.0 | 38.68 | 0.9959 | 0.62 | 1.0 | 0.61 | 1.0 | 0.61 |
| 16 | 1.0 | 47.65 | 0.9938 | 3.49 | 1.0 | 3.53 | 1.0 | 3.53 |
| 17 | 1.0 | 28.94 | 0.9977 | 0.13 | 1.0 | 0.14 | 1.0 | 0.14 |
| 18 | 1.0 | 2.40 | 1.0000 | 0.23 | 1.0 | 0.23 | 1.0 | 0.23 |
| 19 | 1.0 | 2.30 | 1.0000 | 0.93 | 1.0 | 0.93 | 1.0 | 0.93 |
| 20 | 1.0 | 11.45 | 0.9996 | 2.90 | 1.0 | 2.91 | 1.0 | 2.91 |
| 21 | 1.0 | 38.15 | 0.9960 | 0.16 | 1.0 | 0.16 | 1.0 | 0.16 |
| 22 | 1.0 | 153.20 | 0.9370 | 1.58 | 1.0 | 1.63 | 1.0 | 1.63 |
| 23 | 1.0 | 15286.00 | 0.0000 | 16952.00 | 0.0 | 16954.00 | 0.0 | 16954.00 |
| 24 | 1.0 | 4.74 | 0.9999 | 1.44 | 1.0 | 1.44 | 1.0 | 1.44 |
| 25 | 1.0 | 167.63 | 0.9248 | 0.23 | 1.0 | 0.23 | 1.0 | 0.23 |
| 26 | 1.0 | 12.52 | 0.9996 | 0.44 | 1.0 | 0.44 | 1.0 | 0.44 |
| 27 | 1.0 | 70.32 | 0.9865 | 0.72 | 1.0 | 0.74 | 1.0 | 0.74 |
| 28 | 1.0 | 35.97 | 0.9965 | 0.39 | 1.0 | 0.39 | 1.0 | 0.39 |
| 29 | 1.0 | 65.10 | 0.9885 | 0.10 | 1.0 | 0.11 | 1.0 | 0.11 |
| 30 | 1.0 | 4175.90 | 0.0000 | 4985.00 | 0.0 | 4986.50 | 0.0 | 4986.50 |
| 31 | 1.0 | 5624.40 | 0.0000 | 6306.40 | 0.0 | 6303.80 | 0.0 | 6303.80 |
| $\Psi(\mathbf{z})$ | 26533.69 | | 26.15 | | 26.64 | | 26.64 | |

accuracy of the estimated transformation parameters are shown in Fig. 4.8-d. It can be seen that most of the feature points from both images are accurately registered, i.e., the crosses coincide with the squares. The obtained registration error ($D_M$ and $D_{STD}$) are computed as 0.72 and 0.45 pixel, respectively, and this illustrates the registration accuracy. Further, visual inspection shows the continuity of image texture at the border of the overlaid image.

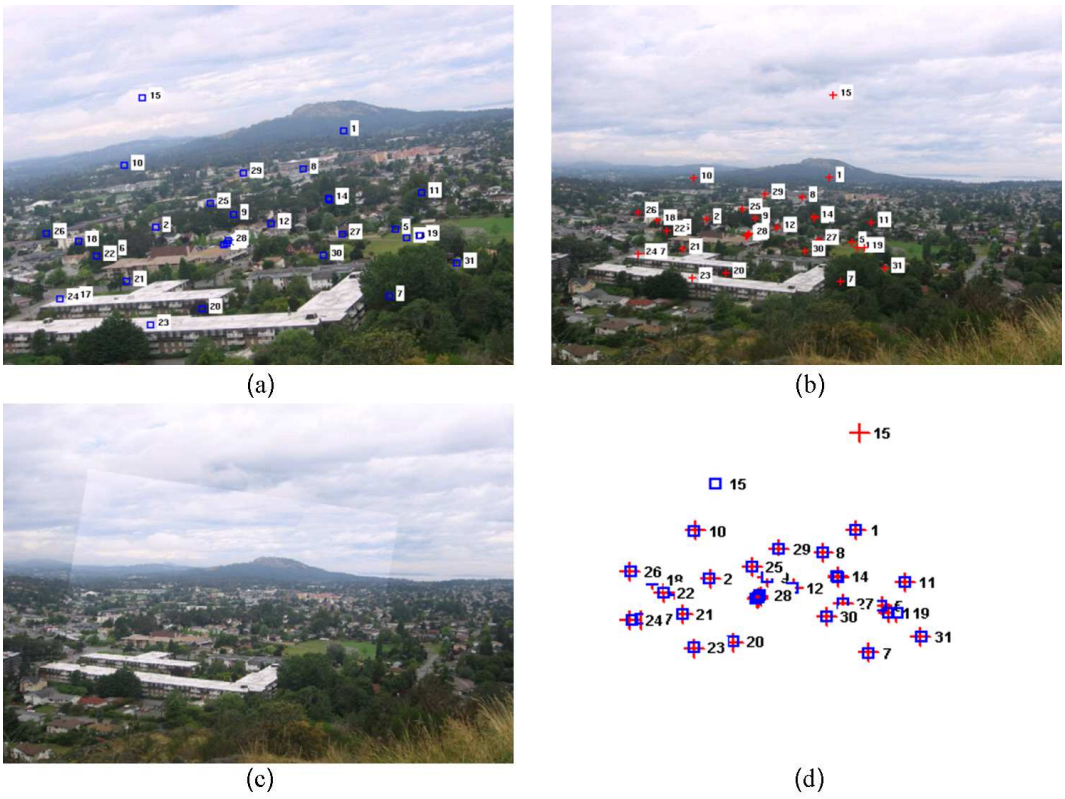



Figure 4.8: Example 2: Registration of two images of an outdoor view taken using a digital camera. (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

- **Example 3**: Two $498 \times 498$ images of an outdoor view of a building were taken using a digital camera. The position of the camera was different for each image and thus the images have only partial overlap, but there was not any zooming. In order to evaluate the robustness of the proposed algorithm to the shearing operation, the target image is further transformed and distorted and the proposed algorithm will be used to register the two images. The reference image is shown in Fig. 4.9-a and the following four degraded images will be considered as the target images.

  ▷ I. The target image is degraded by Gaussian noise of $zero-$mean and standard deviation of 36, and shearing transformation in the x-direction with shearing factor 0.2. It is shown in Fig. 4.9(I)-b

  ▷ II. The target image is degraded by 'Salt & pepper' noise of density 0.06 (affects approximately 6% of the image area) and shearing transformation in both x- and y-directions with shearing factors 0.1 and 0.1, respectively. It is shown in Fig. 4.9(II)-b

  ▷ III. The degraded image is degraded by changing the brightness by $-35$ and shearing transformation in the y-direction with shearing factor $-0.2$ was applied. It is shown in Fig. 4.9(III)-b.

  The proposed technique is used with $s_{p_i} = 1$ for both images since there is no scale difference between the two images. Results of image registration using the proposed algorithm for all cases of example 3 are shown in Fig. 4.9. In this figure, the first and the second column contain the images to be registered, the third column illustrates the correspondence between the feature points and the accuracy of the estimated transformation parameters, and the forth column shows the registered images overlaid over the corresponding reference images. It can be seen in Fig 4.9-c that the crosses (feature points of the reference image)

are matched with the squares (feature points of the second image), except for outliers which have been eliminated by the iterative weighted least squares algorithm by assigning weights to them which are close or equal to zero. The quality of registration is evaluated by computing the registration errors ($D_M$ and $D_{STD}$) listed in Table 4.2. The quality of registration is further confirmed with visual inspection of the borders between the overlaid images.

Table 4.2: Registration errors for Example 3.

| Example | 3-I | 3-II | 3-III |
|---|---|---|---|
| Mean $D_M$ (pixels) | 2.4756 | 1.2961 | 1.9788 |
| Standard deviation $D_{STD}$ (pixels) | 1.1951 | 0.8736 | 1.5623 |

- **Example 4**: Here the proposed technique is applied to different sensor images. In Fig. 4.10, two different sensor images (Urban SPOT band 3 (08/08/95) and TM band 4 (06/07/94)) [2] are shown in the upper row. Difference of the content between the two images can be easily noticed. Fig. 4.10-c shows the registered image overlaid on the reference image and Fig. 4.10-d illustrates the correspondence between the feature points and the accuracy of the estimated transformation parameters. It shows that most of the feature points from both images are accurately registered, where the crosses matched with the squares. The obtained registration error ($D_M$ and $D_{STD}$) are computed as 0.70 and 0.50 pixel, respectively, and this illustrates the registration accuracy. Further, visual inspection at the borders between the registered images shows the continuity of the coastlines and confirms the quality of registration.

[2] Available at http://nayana.ece.ucsb.edu/registration/examples.shtml

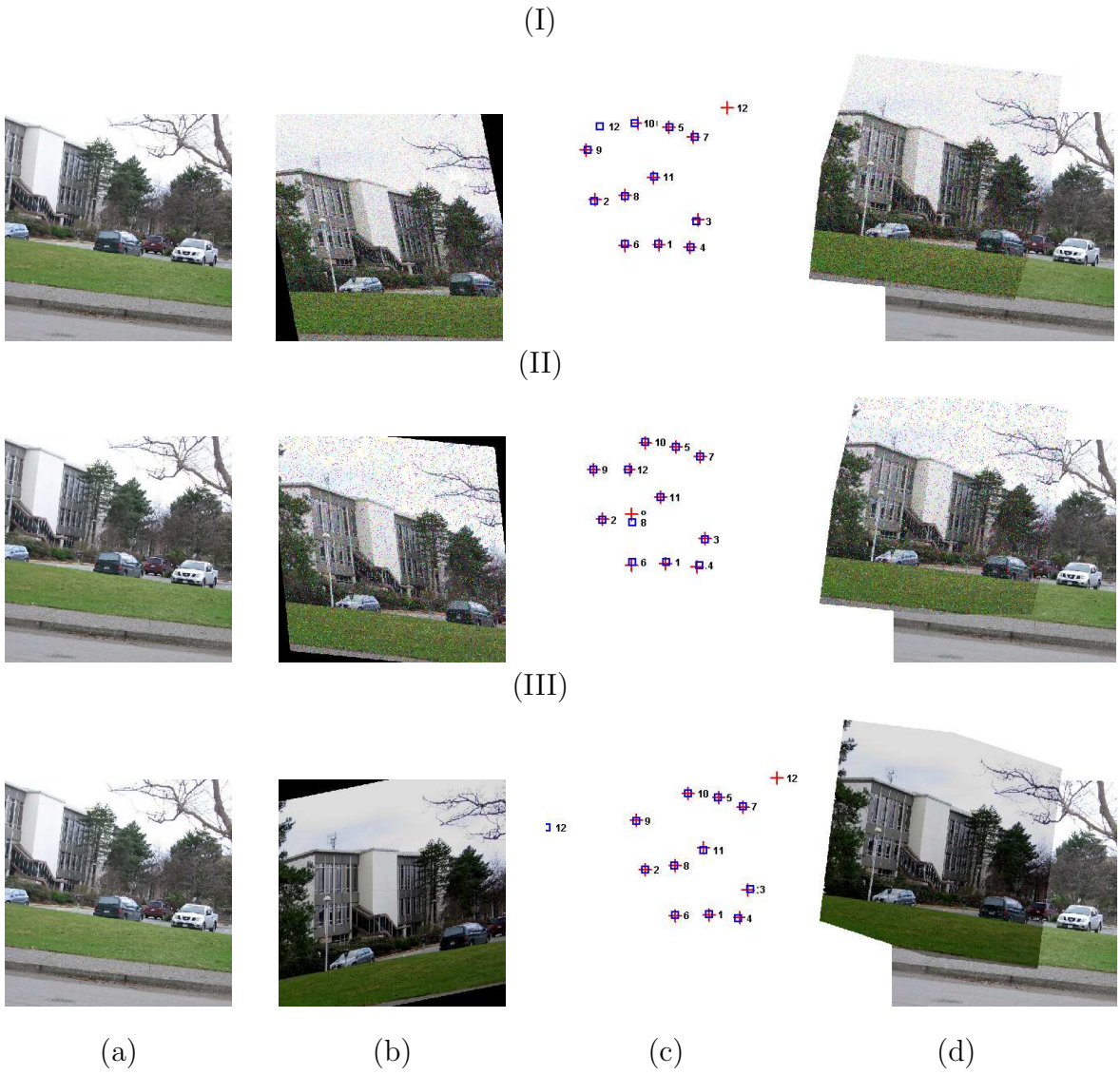



Figure 4.9: Example 3: Registration of images having additional shearing distortions. (a) Reference images; (b) Target images; (c) Correspondence between the feature points of the reference images (represented by crosses) and the feature points of registered (transformed) images (represented by squares); (d) The registered images overlaid on the corresponding reference images.

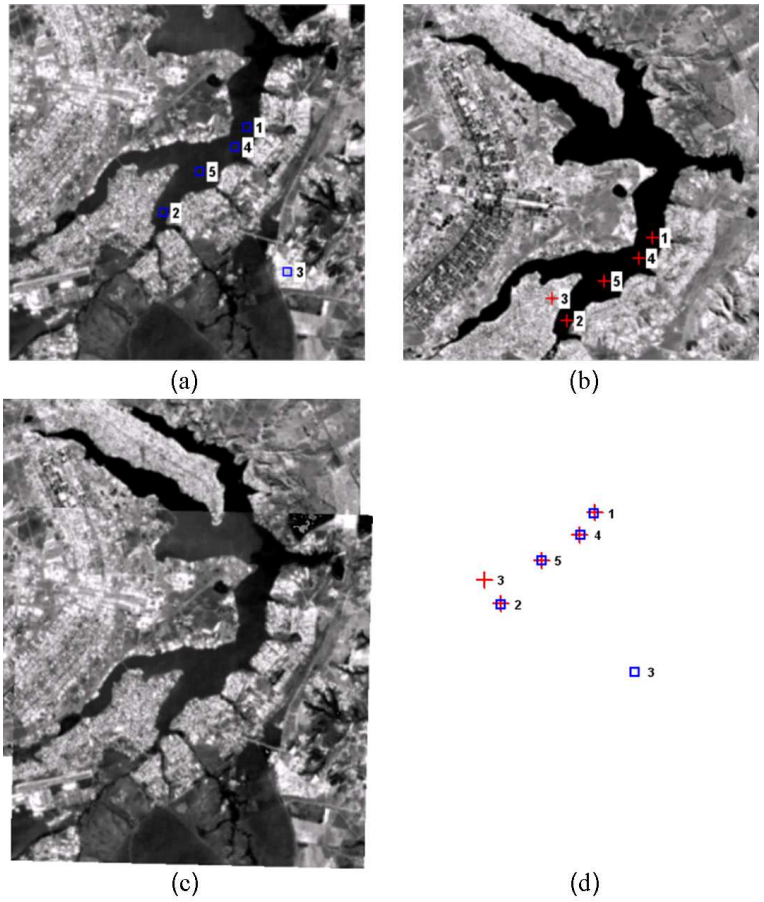



Figure 4.10: Example 4: Registration of two different sensor images (Urban SPOT band 3 (08/08/95) and TM band 4 (06/07/94)). (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

- **Example 5**: Fig. 4.11 shows another example of registering two high resolution Ikonos optical images of UCSB, University of California, Santa Barbara site [3].The upper row shows the two images. Fig. 4.11-c shows the registered image overlaid on the reference image. The correspondence between the feature points and the accuracy of the estimated transformation parameters are shown in Fig. 4.11-d. It can be seen that the feature points from both images are accurately registered, where the crosses matched with the squares, except of the outliers which have been correctly eliminated. The obtained registration error ($D_M$ and $D_{STD}$) are computed as 0.49 and 0.23 pixel, respectively which illustrates the registration accuracy. Further, the continuity of the coastlines, the roads and buildings can be easily noticed.

- **Example 6**: Another example of two images which have scale, rotation and translation changes is shown in Fig. 4.12. The upper row shows the two images [4] and Fig. 4.12-c shows the registered image overlaid on the reference image. The correspondence between feature points and the accuracy of the estimated transformation parameters are shown in Fig. 4.12-d. The obtained registration error ($D_M$ and $D_{STD}$) are computed as 0.51 and 0.22 pixel, respectively, and this illustrates the quality of registration. Further, the continuity of the mast and boat edges on the border of the overlaid image can be easily noticed.

### 4.3.2 Comparison with Other Techniques

The experimental results presented in the previous section showed that the proposed technique performs very well in a variety of affine transformations and can give good results despite image degradations such as noise, blurring and other image processing

[3]Available at http://nayana.ece.ucsb.edu/registration/examples.shtml

[4]Available at http://www.robots.ox.ac.uk/ vgg/research/affine/

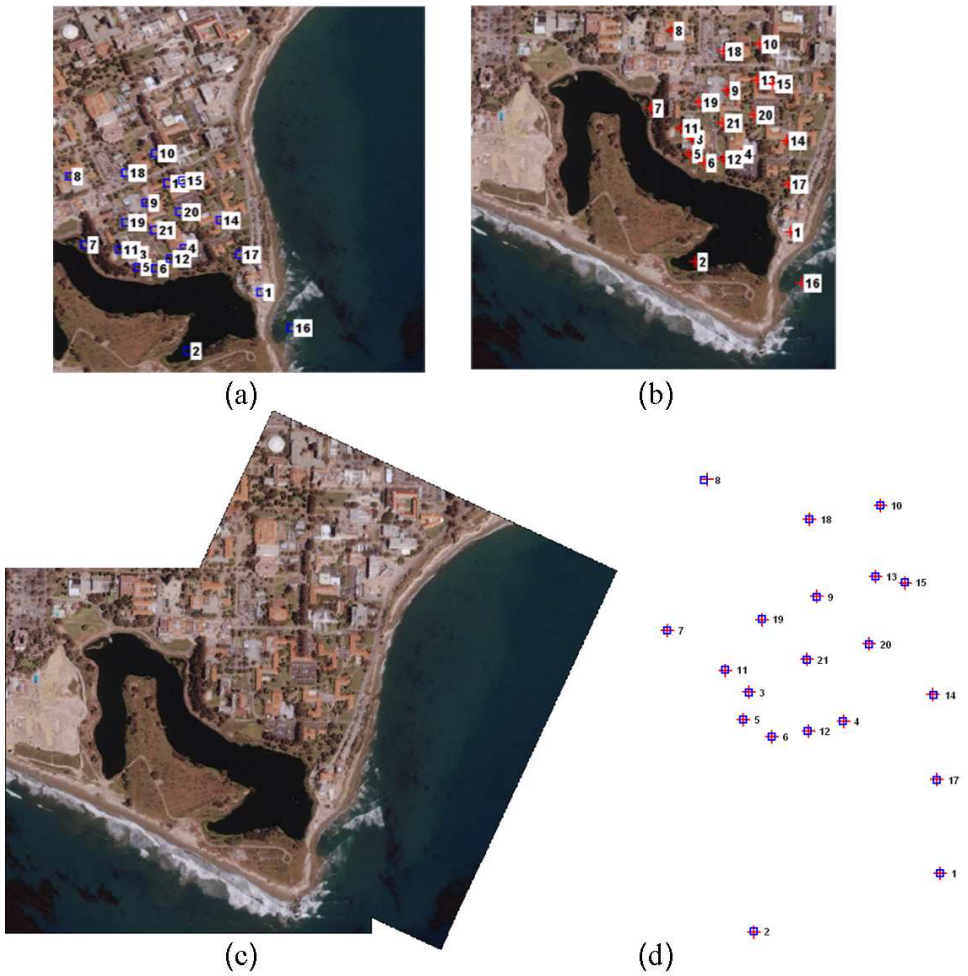



Figure 4.11: Example 5: Registration of two high resolution Ikonos optical images of UCSB, University of California, Santa Barbara site. (a) and (b) The two images with the paired feature points superimposed on each image; (c) The registered image is overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

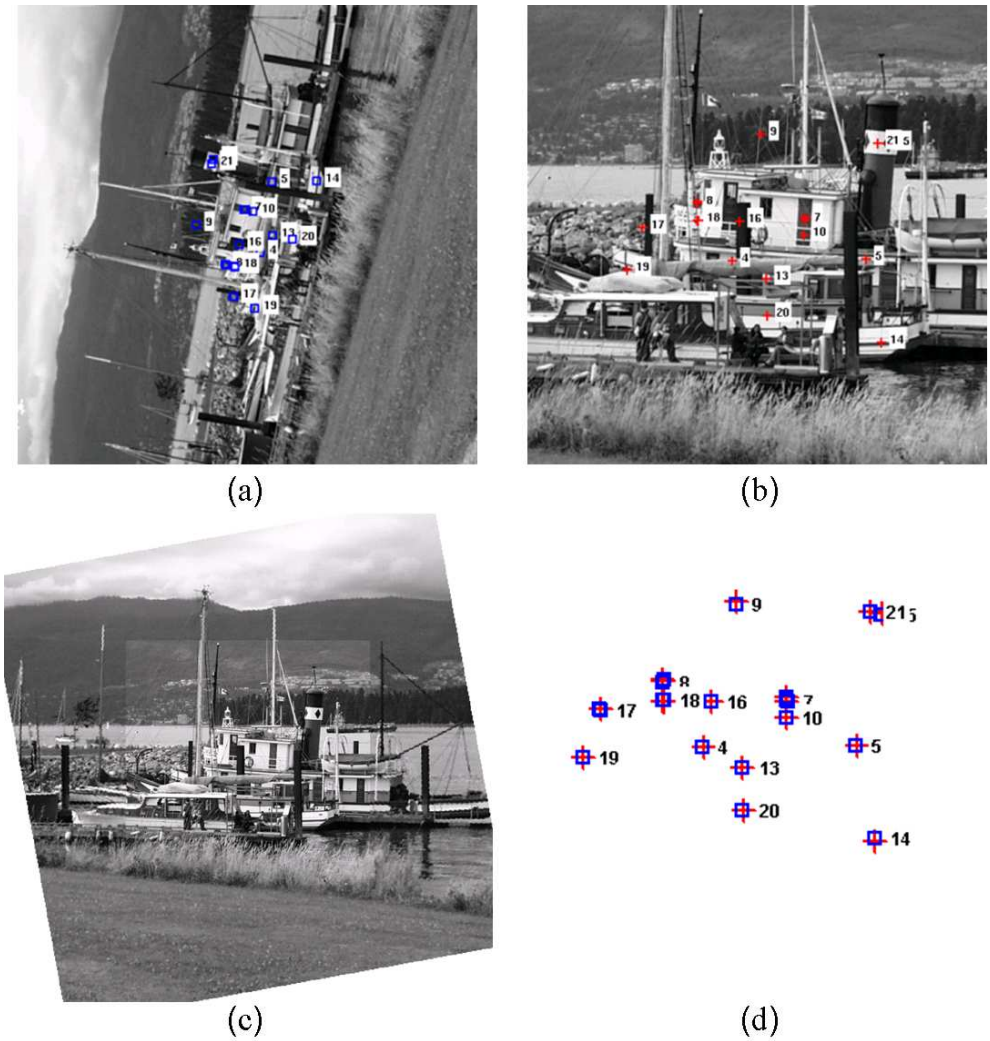



Figure 4.12: Example 6: Registration of two photographs having different scale of an outdoor scene. (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

operations. One interesting question to be considered here is how is this performance compared to other registration techniques presented in the literature. A comparison with the methods in [16], [33] and [19] will be discussed here.

- For a comparison with the technique in [16], both methods are applied to register image pairs shown in the upper rows in Fig. 4.13 and 4.14 [5]. The images used in Fig. 4.13 are Landsat Thematic Mapper (TM) Band 0 and Band 8 images while the images used in Fig. 4.14 are aerial images of the Mojave desert taken by a balloon flying over the area. For comparing the performance of the two methods, the root mean squared error RMSE between the reference feature points and their registered correspondents is used as a measure as it was done in [16]. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{\widetilde{K}} \sum_{i=1}^{\widetilde{K}} (a_{11}x'_{1i} + a_{12}x'_{2i} + t_{x1} - x_{1i})^2 + (a_{21}x'_{1i} + a_{22}x'_{2i} + t_{x2} - x_{2i})^2}, \tag{4.23}$$

  where $\widetilde{K}$ is the total number of point pairs. The RMSE obtained by manual registration [16] , by using the method in [16] and by using our proposed method is shown in Table 4.3. As it can be seen in Table 4.3, the method proposed here gives lower RMSE values than the values obtained in [16]. The registration results obtained using the proposed method can be seen in the lower-left figures that shows the registered images overlaid on the reference ones in Fig. 4.13 and 4.14.

- The proposed registration method is also compared with a publicly available registration tool imREG [6] [33]. The registration method proposed in this chapter and the one in [33] were applied to register image pairs from the test

[5] Available at http://vision.ece.ucsb.edu/registration/satellite/testimag/index.htm

[6] http://nayana.ece.ucsb.edu/registration/

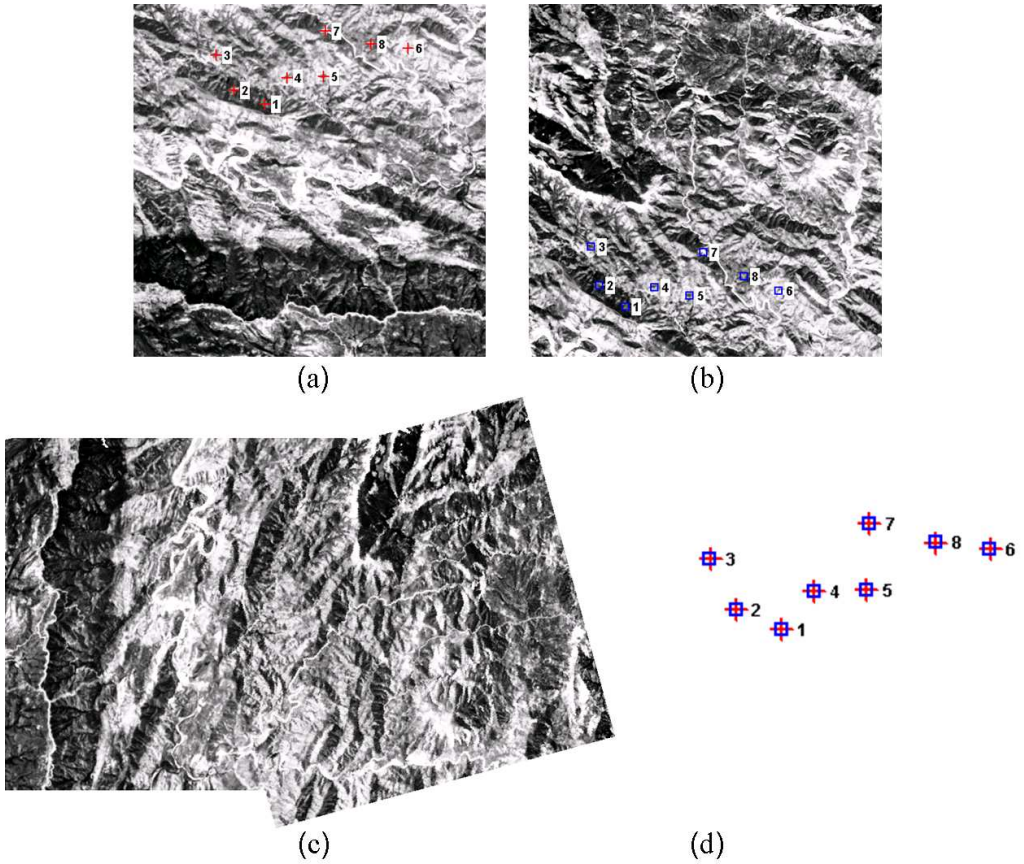



Figure 4.13: Registration of a pair of Landsat Thematic Mapper (TM) Band 0 and Band 8 images: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

Table 4.3: RMSE (in pixels) for images in Fig. 4.13 and 4.14.

| Image | Manual [16] | Contour-based method [16] | Proposed method |
|---|---|---|---|
| Landsat TM images | 1.68 | 0.61 | 0.14 |
| Mojave desert images | 1.82 | 0.69 | 0.32 |

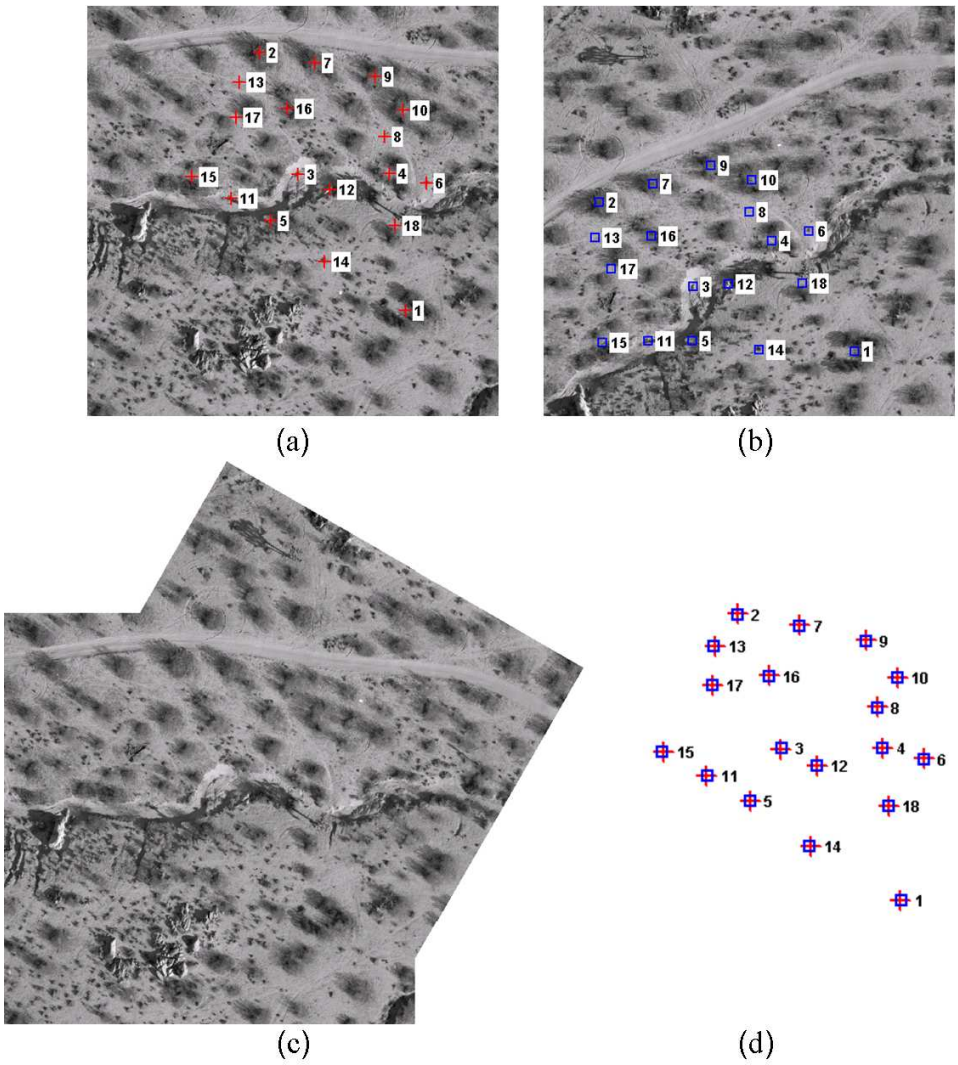



Figure 4.14: Registration of a pair of aerial images of the Mojave desert: (a) and (b) The two images with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

images library of the imREG registration tool, in addition to other image pairs. The image pairs used are shown in Fig. 4.15.

The RMSE, Eqn. (4.23), for these examples obtained using the imREG tool and the proposed method can be found in Table 4.4. It can be seen in this table that the proposed registration algorithm yields lower registration errors than the imREG tool. The imREG tool failed in registering images of different scales, as in Fig. 4.15-Example 4 and 6, but it was successful in registering multitemporal images (images that were taken at different times), as in Fig. 4.15-Example 8.

- In order to compare the performance of the proposed method for image registration of images of different scales, the scale-invariant feature detector SIFT [19] was used. In this test, the first and the second steps of the proposed algorithm were replaced by the feature detection and descriptors matching processes used in the SIFT method [19]. The image pairs shown in Fig. 4.16 were used as test images for comparing the registration accuracy using the SIFT method to that of the proposed method. The obtained registration errors ($D_M$,$D_{STD}$) for these image pairs using the proposed method and using the SIFT method are found in Table 4.5. It can be noticed from comparing the obtained registration errors that the proposed method yields comparable results as the SIFT method.

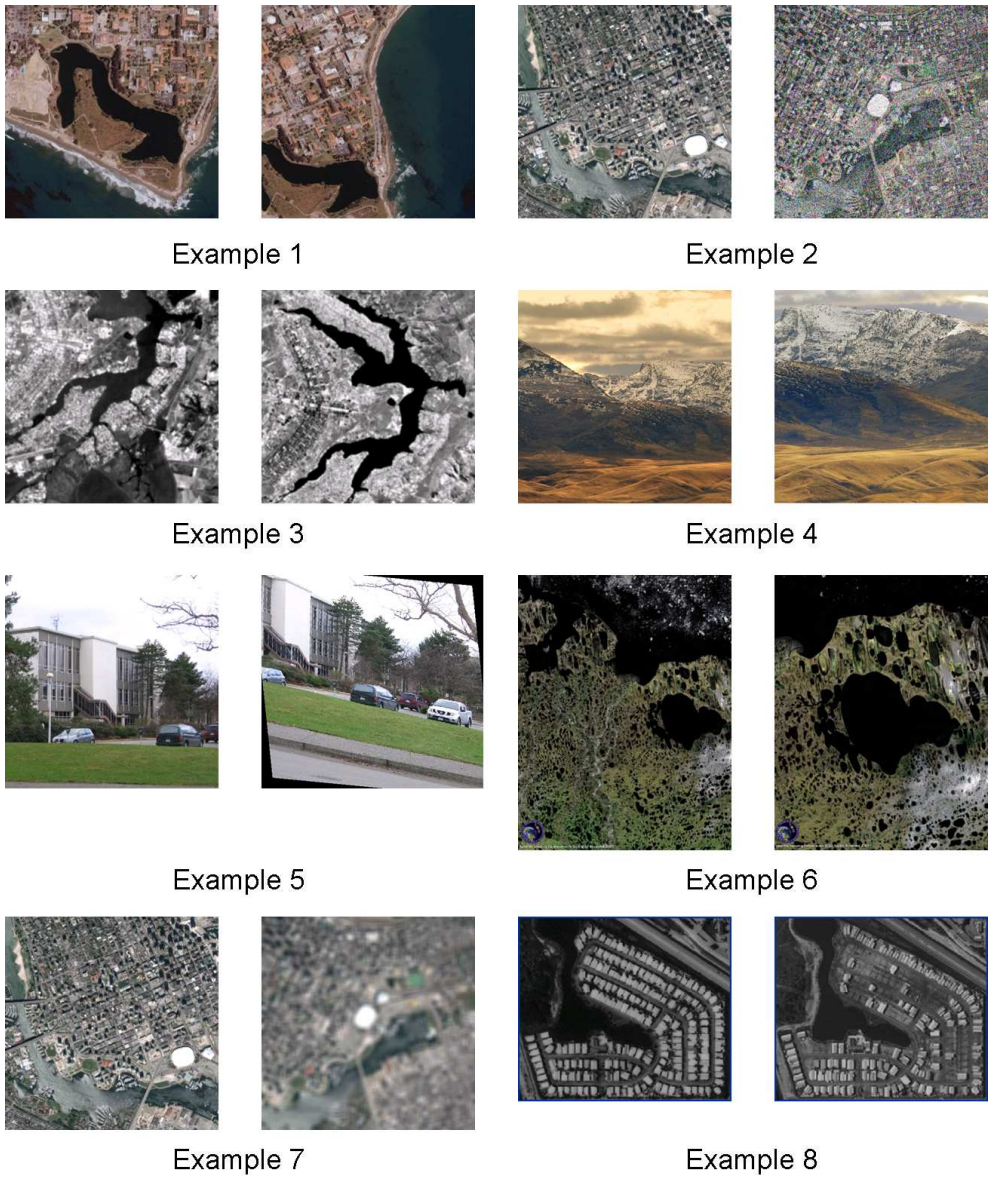



Figure 4.15: Examples for comparison with imREG tool.

Table 4.4: RMSE (in pixels) of comparison examples (in Fig. 4.15) with imREG tool [33].

| Image | imREG tool [33] | Proposed method |
|---|---|---|
| Example 1 | 0.77 | 0.45 |
| Example 2 | 1.80 | 0.74 |
| Example 3 | 2.45 | 1.60 |
| Example 4 | Failed | 1.68 |
| Example 5 | 1.74 | 0.35 |
| Example 6 | Failed | 0.84 |
| Example 7 | Failed | 0.91 |
| Example 8 | 1.03 | Failed |

Table 4.5: Registration errors ($D_M$,$D_{STD}$) of comparison examples (in Fig. 4.16) with SIFT method [19].

| | SIFT method [19] | Proposed method |
|---|---|---|
| Example 1 | $(0.40, 0.19)$ | $(0.86, 0.46)$ |
| Example 2 | $(0.74, 0.46)$ | $(0.72, 0.44)$ |
| Example 3 | $(1.25, 1.01)$ | $(0.51, 0.22)$ |
| Example 4 | $(0.95, 0.51)$ | $(0.66, 0.33)$ |
| Example 5 | $(1.08, 1.58)$ | $(0.74, 0.45)$ |
| Example 6 | $(0.89, 0.64)$ | $(1.01, 0.92)$ |
| Example 7 | $(0.28, 0.26)$ | $(0.30, 0.12)$ |
| Example 8 | $(0.40, .30)$ | $(0.26, 0.11)$ |
| Example 9 | $(0.32, 0.42)$ | $(0.94, 0.66)$ |
| Example 10 | $(0.28, 0.19)$ | $(0.63, 0.84)$ |

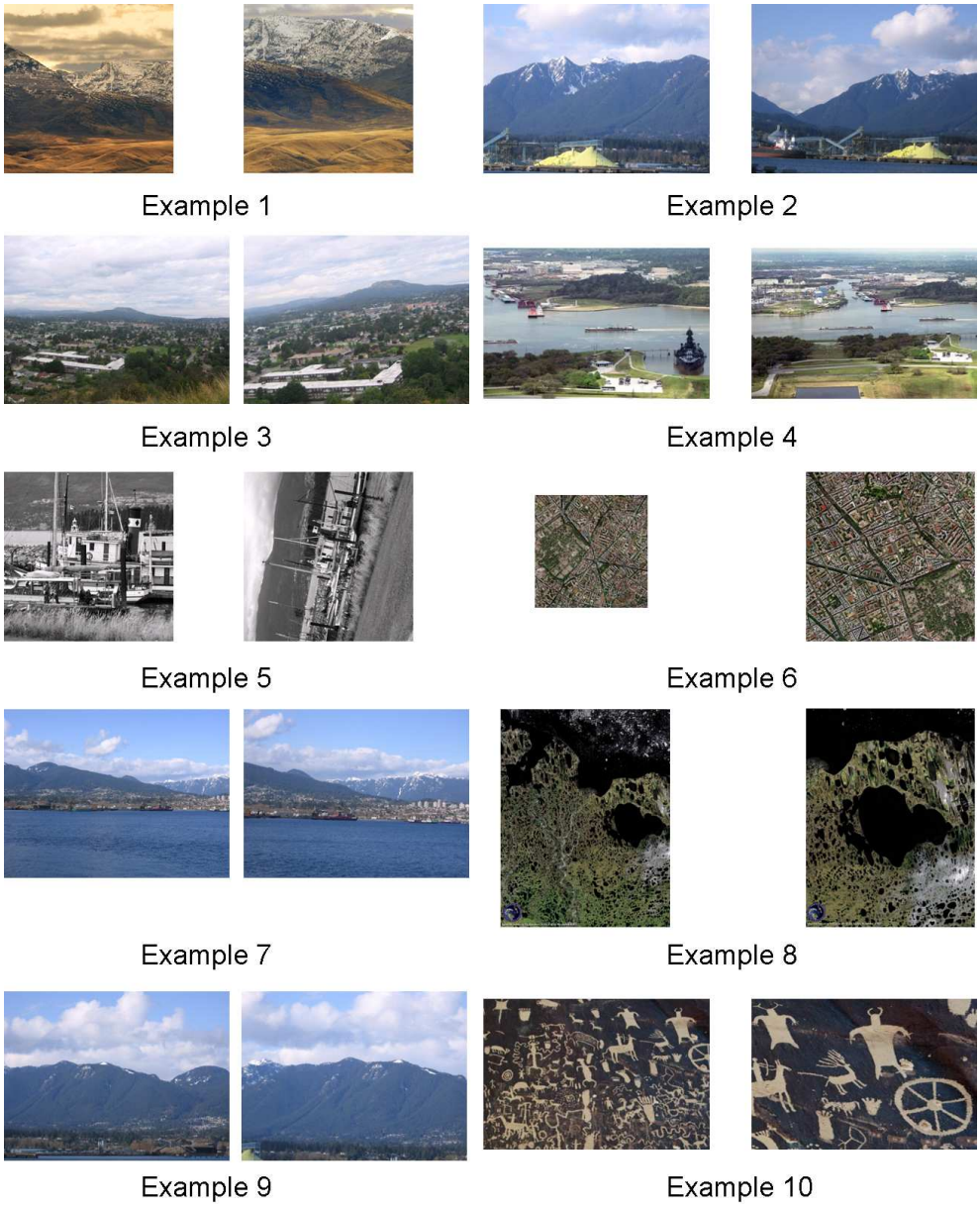



Figure 4.16: Examples for comparison with SIFT method.

## 4.4 Conclusions

In this chapter, an algorithm for registration of images with affine geometric distortions was discussed. The algorithm is based on three main steps: extracting feature points in images, obtaining the correspondence between the feature points of the images and estimating the transformation parameters which map the target image to the reference one using an iterative weighted lease squares algorithm.

The performance of the proposed algorithm was evaluated using several types of images and degradations. Based on the obtained results, it can be concluded that:

- The proposed algorithm gives excellent registration accuracy using several types of images, even in cases with partial overlap between images.
- In terms of robustness, the obtained experimental results indicate that the proposed algorithm is robust against many image degradations such as blurring, noise contamination, brightness change, etc.
- Experimental results indicate that the proposed algorithm can handle images of different scales effectively.

# Chapter 5

# A Geometric Normalization Technique and its Application in Image Watermarking

This chapter introduces a technique for image geometric normalization based on the geometric properties of image feature points. The proposed technique consists of two main steps, feature point extraction and normalization parameters estimation. A technique for image watermarking with robustness to geometric distortions will be also introduced as an example of the possible applications of the presented geometric normalization technique.

This chapter is organized as follows. Section 5.1 discusses the proposed image normalization technique: its objectives and requirements. Section 5.2 describes the proposed normalization technique in detail. In Section 5.3, experimental results are presented and the performance of the proposed technique is discussed. Section 5.4 describes the proposed watermarking technique and evaluates its performance. Finally, Section 5.5 summarizes the chapter.

## 5.1 Introduction

Image normalization techniques have been used as a preprocessing step in many applications such as pattern recognition and classification [1], image retrieval [2], and image watermarking [65]. Image normalization is used in such applications with the purpose of representing objects, patterns (or the entire image) regardless of changes in their orientation, size, or position [1].

### 5.1.1 Objectives

This chapter introduces a technique for image geometric normalization based on the geometric properties of image feature points, subject to the following objectives:

- The normalization technique should be robust to image degradations such as image blurring, brightness change and noise contamination. In addition, cropping parts of the image should not affect the obtained normalization results.
- As an image normalization technique is usually used as a preprocessing step in certain applications, it should be as simple and fast as possible.
- The normalization technique should be suitable to be employed in different applications.

### 5.1.2 Geometric Distortions Considered in the Normalization Process

The geometric distortions considered here for geometric normalization are rotation, scaling and translation, and can be defined as

$$\mathbf{p}_n = T(t_{x_1}, t_{x_2}) + sR(\theta)\mathbf{p}, \tag{5.1}$$

where $\mathbf{p}=(p_{x_1}, p_{x_2})$ is a point of the image $I$, which is mapped to a point $\mathbf{p}_n=(p_{x_1 n}, p_{x_2 y_n})$ of the normalized image $I_n$, $T$ represents the translation operation,

and $R$ represents the rotation operation. In detail, such transformation equation can be rewritten as

$$\begin{pmatrix} p_{x_1 n} \\ p_{x_2 n} \end{pmatrix} = \begin{pmatrix} t_{x_1} \\ t_{x_2} \end{pmatrix} + s \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} p_{x_1} \\ p_{x_2} \end{pmatrix}, \tag{5.2}$$

where $t_{x_1}$ and $t_{x_2}$ represent translation parameters in $x$ and $y$ directions, respectively, $s$ is the scaling parameter and $\theta$ is the rotation angle.

## 5.2 The Proposed Geometric Normalization

The proposed normalization technique consists of two main steps: feature point extraction and normalization parameters estimation. The steps of the proposed technique are shown schematically in Fig. 5.1. The two main steps of the technique are described in detail in the following subsections.

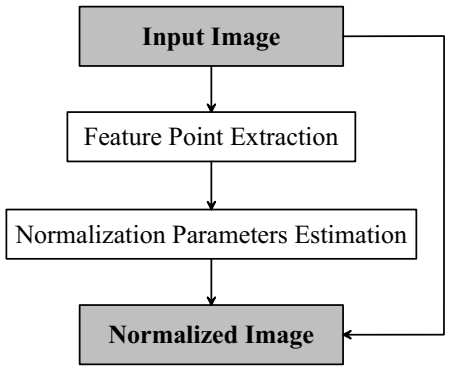



Figure 5.1: Block diagram of the proposed normalization technique.

### 5.2.1 Feature Point Extraction

A set of feature points is extracted from the given image using the feature point extractor presented in Chapter 3.

Assuming a square input image of size $N_i \times N_i$ with an active image area $A_i$, the feature point extractor is used with $s_{p_i} = \frac{N_n^2}{A_i}$. This compensates for scale difference between the input image and a standard size of $N_n \times N_n$. The purpose of this initial scaling is to make the scale relationship between the Mexican-hat wavelet size and the image features for all images to be normalized consistent. This is an essential step even though $s_{p_i}$ may not be precise due to degradations and possible cropping parts of the image. Fortunately, large cropping is not common as it destroys the original image. Moreover, small imprecisions of the Mexican-hat wavelets and neighborhoods sizes do not affect most of the extracted feature points either. This is demonstrated by examples in Section 5.3.

The images used in the subsequent examples have the size of $512 \times 512$ and, therefore, $N_n$ is used with value of 512.

The locations of the obtained feature points are found as $(P_i, i = 1, 2, \ldots, K)$, where $P_i = (p_{ix}, p_{iy})$ are the coordinates of a point $P_i$ and $K$ is the number of feature points.

### 5.2.2 Normalization Parameters Estimation

The locations of the extracted feature points are used here to estimate the parameters for the image normalization process.

1. The obtained feature points are sorted according to the value of the feature detection response $\phi$, in Eqn. 3.3, and the three points with the highest values are taken as the basis of the normalization process. Assume that these points $P_0$, $P_1$ and $P_2$ have coordinates $(p_{0x_1}, p_{0x_2})$, $(p_{1x_1}, p_{1x_2})$ and $(p_{2x_1}, p_{2x_2})$, respectively.

2. The points $P_0$, $P_1$ and $P_2$ form a triangle, which can be used to estimate the normalization parameters. The geometric attributes used are as follows:

   (a) *Incenter* $P_c$ of the triangle $P_0P_1P_2$. $P_c$ is the center of the circle inscribed in the triangle and can be obtained as the intersection of the three angle bisectors. The coordinates of $P_c$ can be obtained as

   $$P_c = (x_{1c}, x_{2c}) = (\frac{\ell_0 p_{0_{x_1}} + \ell_1 p_{1_{x_1}} + \ell_2 p_{2_{x_1}}}{\ell_0 + \ell_1 + \ell_2}, \frac{\ell_0 p_{0_{x_2}} + \ell_1 p_{1_{x_2}} + \ell_2 p_{2_{x_2}}}{\ell_0 + \ell_1 + \ell_2}), \tag{5.3}$$

   where the opposite sides of the triangle vertices ($P_0$, $P_1$ and $P_2$) have lengths $\ell_0, \ell_1$ and $\ell_2$, respectively.

   (b) *The vertex* $P_m$ *of the largest angle of the triangle* $P_0P_1P_2$. The coordinates of the vertex $P_m$ $(p_{m_{x_1}}, p_{m_{x_2}})$ will be the coordinates of one of the three vertices $P_0$, $P_1$ or $P_2$.

   (c) *The length* $\ell_m$ *of the longest side of* $P_0P_1P_2$, which is the length of the triangle side that is the longest among the three sides and it is the opposite side of $P_m$.

3. The translation normalization is achieved by translating the image so that the incenter $P_c$ coincides with the center of a new image $(x_{1cn}, x_{2cn})$ using the translation parameters

   $$(t_{x_1}, t_{x_2})^T = (x_{1cn} - x_{1c}, x_{2cn} - x_{2c})^T \tag{5.4}$$

4. The rotation normalization is achieved by rotating the image so that the positive x-axis of the normalized image coincides with the line $\overline{P_cP_m}$, see Fig. 5.2:

   $$\theta = \tan^{-1} \frac{p_{m_{x_2}} - x_{2c}}{p_{m_{x_1}} - x_{1c}}, \tag{5.5}$$

   where $P_m = (p_{m_{x_1}}, p_{m_{x_2}})$.

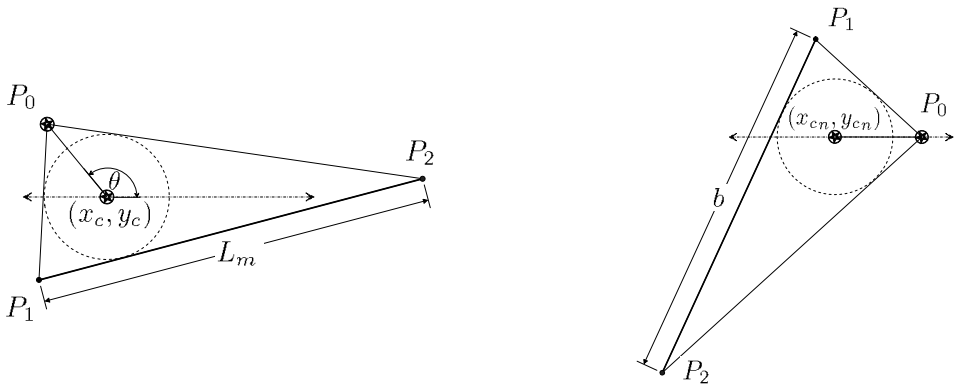
5. For scaling normalization, the scaling factor $s$ is estimated as the ratio between $\ell_m$ and $b$, where $b$ is a predefined normalization length for the longest side of the triangle in the normalized image:

$$s = \frac{b}{\ell_m}, \tag{5.6}$$

where the length $\ell$ between two points $(p_{i_{x_1}}, p_{i_{x_2}})$ and $(p_{j_{x_1}}, p_{j_{x_2}})$ can be obtained as

$$\ell = \sqrt{(p_{i_{x_1}} - p_{j_{x_1}})^2 + (p_{i_{x_2}} - p_{j_{x_2}})^2} \tag{5.7}$$

6. Finally, the estimated translation parameters $(t_{x_1}, t_{x_2})^T$, the rotation angle $\theta$, and the scaling factor $s$ are used in transforming the input image $I$ to the normalized form $I_n$, using Eqn. 5.2.




(a) (b)

Figure 5.2: Geometric normalization using the feature points: (a) Estimating the normalization parameters using the three points with the highest detection response values, (b) Transforming the feature points using the estimated normalization parameters.

## 5.3 Experimental Results

In this section, the performance of the proposed normalization technique is evaluated for images with different characteristics and for several types of image degradation. In order to evaluate the accuracy and the robustness of the proposed technique, undistorted original images and some distorted and degraded versions of the same images are normalized using the proposed technique.

- *Normalization examples*:

  In Fig. 5.3, some normalization examples are shown. In this figure, the original and the distorted images are shown in the first column. The second column shows the images with the extracted feature points superimposed on them. The resulting normalized form of each image is shown in the third column. It can be seen in this figure that the obtained normalized images are geometrically identical regardless of the applied distortion or degradation.

- *Evaluating the performance when using two or three feature points in the normalization process*:

  In order to evaluate the performance when using two feature points (as in [66]) or three feature points (as in the proposed technique) in the normalization process, the performance of the proposed technique is compared with that of an earlier technique presented in [66] in terms of normalization errors. In order to quantify the normalization error, the error in the normalization rotation angle $Err_\theta$, the error in the normalization scaling factor $Err_s$ and the error in the normalization translation parameters $Err_t$ are calculated as follows:

  $$Err_\theta = |(\theta_0 - \theta_d) - \theta)|, \tag{5.8}$$

  where $\theta_0$ and $\theta$ are the normalization rotation angles for the original and the distorted images, respectively, and $\theta_d$ is the rotation angle when distorting the

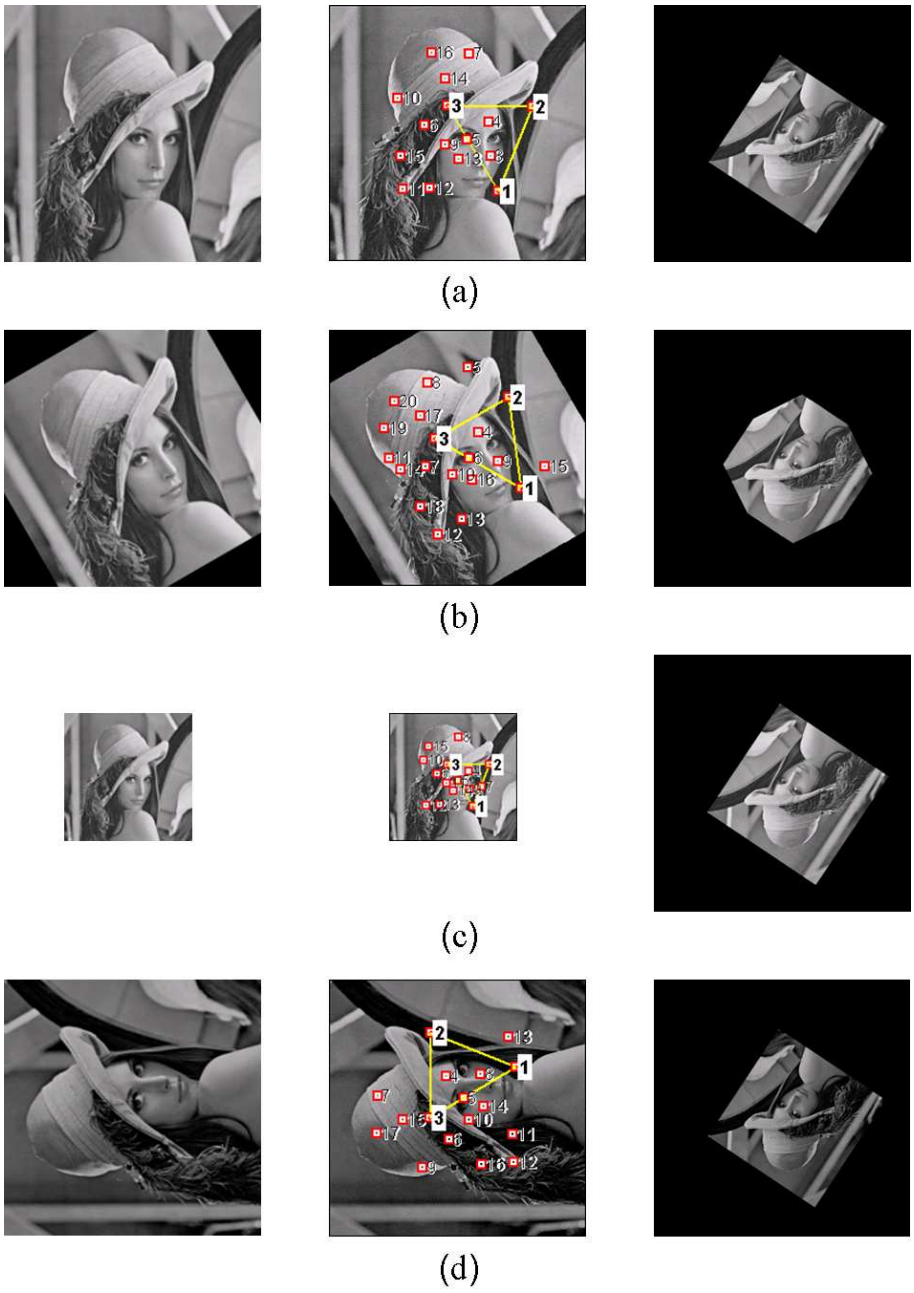



Figure 5.3: Examples of normalizing distorted images: (a) Original, undistorted 'Lena' image, (b) rotation by $30^o$, (c) scaling down by 50%, and (d) brightness change by $(-45)$ and rotation by $90^o$.

original image.

$$Err_s = 100 \times \frac{|(s_0/s) - s_d|}{s_d}\%, \tag{5.9}$$

where $s_0$ and $s$ are the normalization scaling factors for the original and the distorted images, respectively, and $s_d$ is the scaling factor when distorting the original image.

$$Err_t = \sqrt{(x_{1c0} - x_{1ct})^2 + (x_{2c0} - x_{2ct})^2} \tag{5.10}$$

where $(x_{1c0}, x_{2c0})$ is the triangle's incenter in the original image and $(x_{1ct}, x_{2ct})$ is the transformed triangle's incenter in the distorted image, after transforming $(x_{1c}, x_{2c})$ using $-\theta_d$ and $s_d^{-1}$.

Various test images have been used and a set of different distortions/degradations has been applied on each image. The applied distortions include rotation, scaling and cropping parts of the images while the applied degradations include blurring, brightness change, image compression and noise contamination. The normalization errors using both techniques are obtained for four sets of images and the means and the standard deviations of the obtained normalization errors are plotted in Fig. 5.4. In each case, the left graph shows the means of the normalization parameters errors, due to the degradations of the images, while the right graph shows the standard deviation of these errors. It can be seen in that figure that the proposed technique has less normalization errors than the one in [66] and, consequently, leads to more accurate normalization results.

- *Comparison with a moments-based normalization technique*:

  The performance of the proposed technique is also compared with that of a moments-based normalization technique presented in [65]. Various test images have been used and a set of different distortions/degradations has been applied on each image. The test images used are shown in Fig. 5.5. A set of distortions/degradations that includes rotation, scaling, cropping

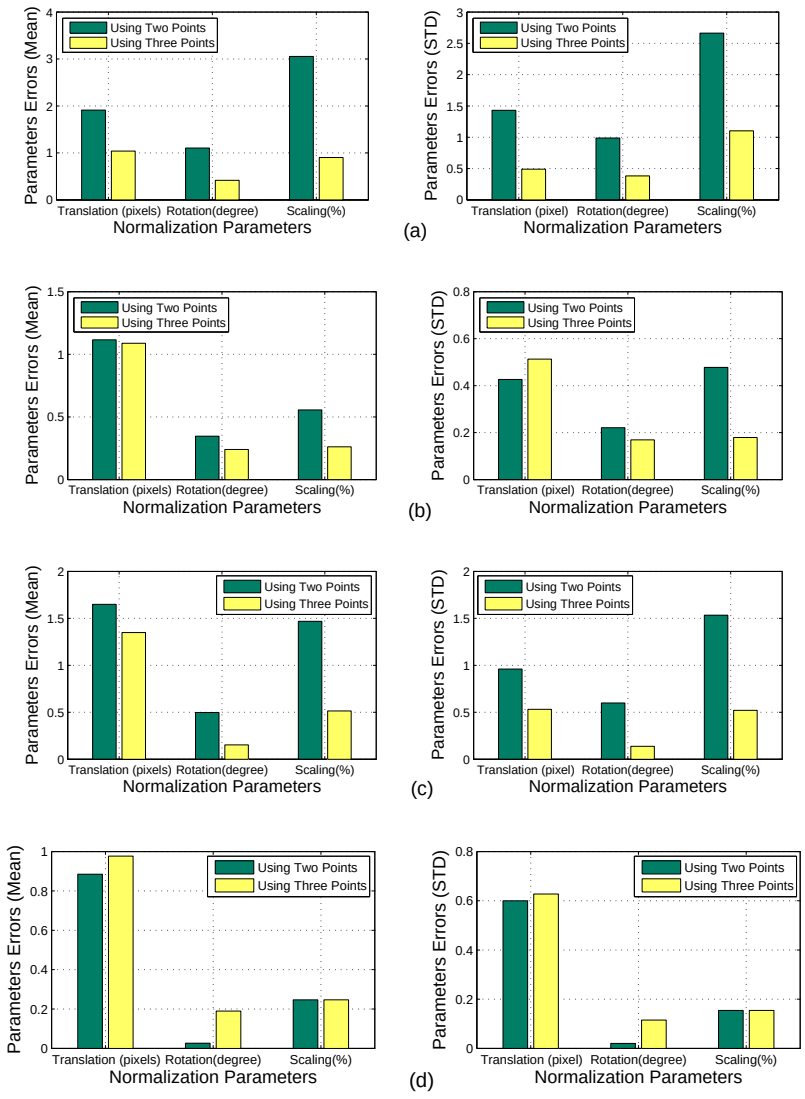



Figure 5.4: Evaluating the accuracy of the estimated normalization parameters when using two or three feature points in estimating the normalization parameters: (a)'Barbara' image , (b) 'Bridge' image, (c) 'Man' image, and (d) 'Airplane' image.

Figure 5.5: The image-set used in evaluating the accuracy of the estimated normalization parameters.

parts of the images, blurring, brightness change, image compression and noise contamination are applied on the images. The normalized forms of the images are obtained and the difference-images between the normalized forms of the original images and the normalized forms of the distorted images are used to assess the accuracy of the proposed technique visually. A completely blank difference-image implies perfectly aligned images and, therefore, a precise normalization. However, in some cases due to additional degradations, the difference-image might contain some non-zero parts even though the result of the normalization is perfect. If, for example, the original image has been changed due to cropping, brightness change, noise addition, etc., then the difference-image will contain some non-zero elements.

In Fig. 5.6 and 5.7, some examples of the obtained difference-images using the proposed technique and using a moment-based method [65] for different distortions/degradations are shown. In each example, the input image is shown in the first column; the resulting difference-image for the proposed technique is shown in the second column; and the resulting difference-image for the moments-based technique [65] is shown in the third column.

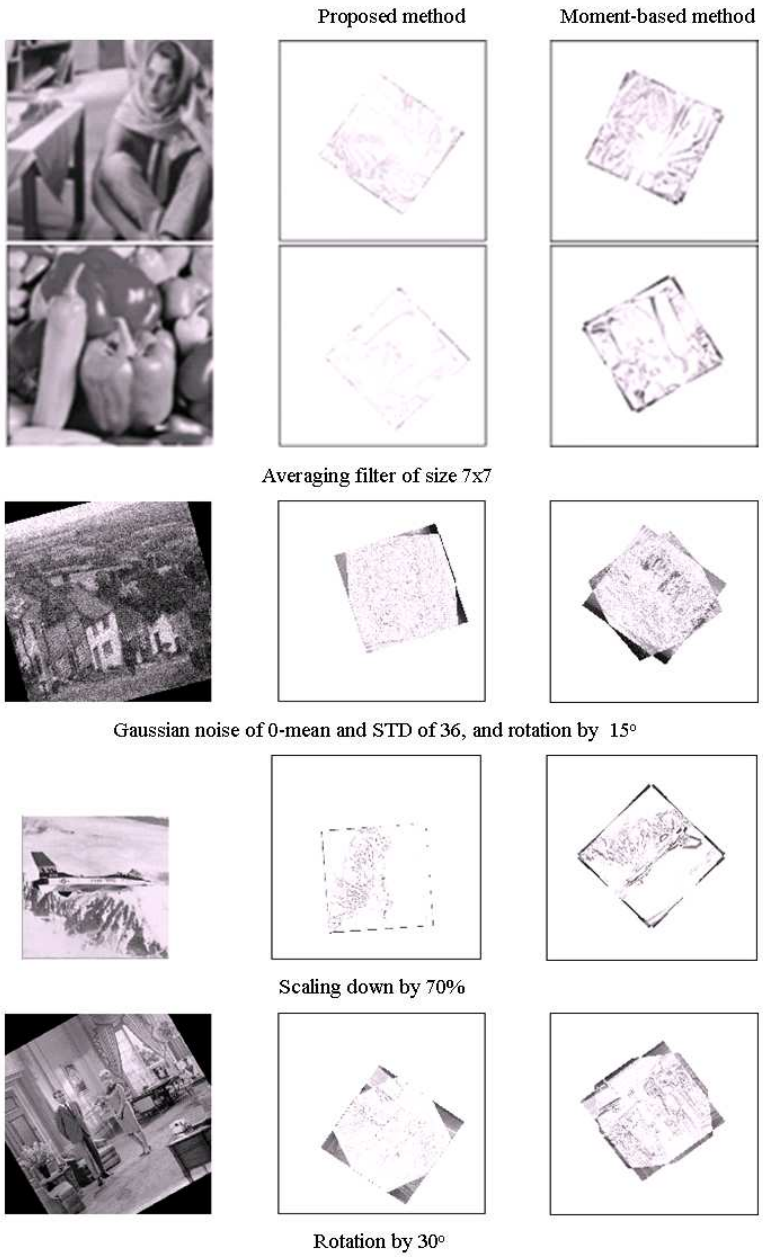



Figure 5.6: Evaluating the obtained normalization results: the obtained difference-images using the proposed technique and using the moment-based method in [65] for different distortions/degradations (blurring, Gaussian noise contamination, scaling and rotation).

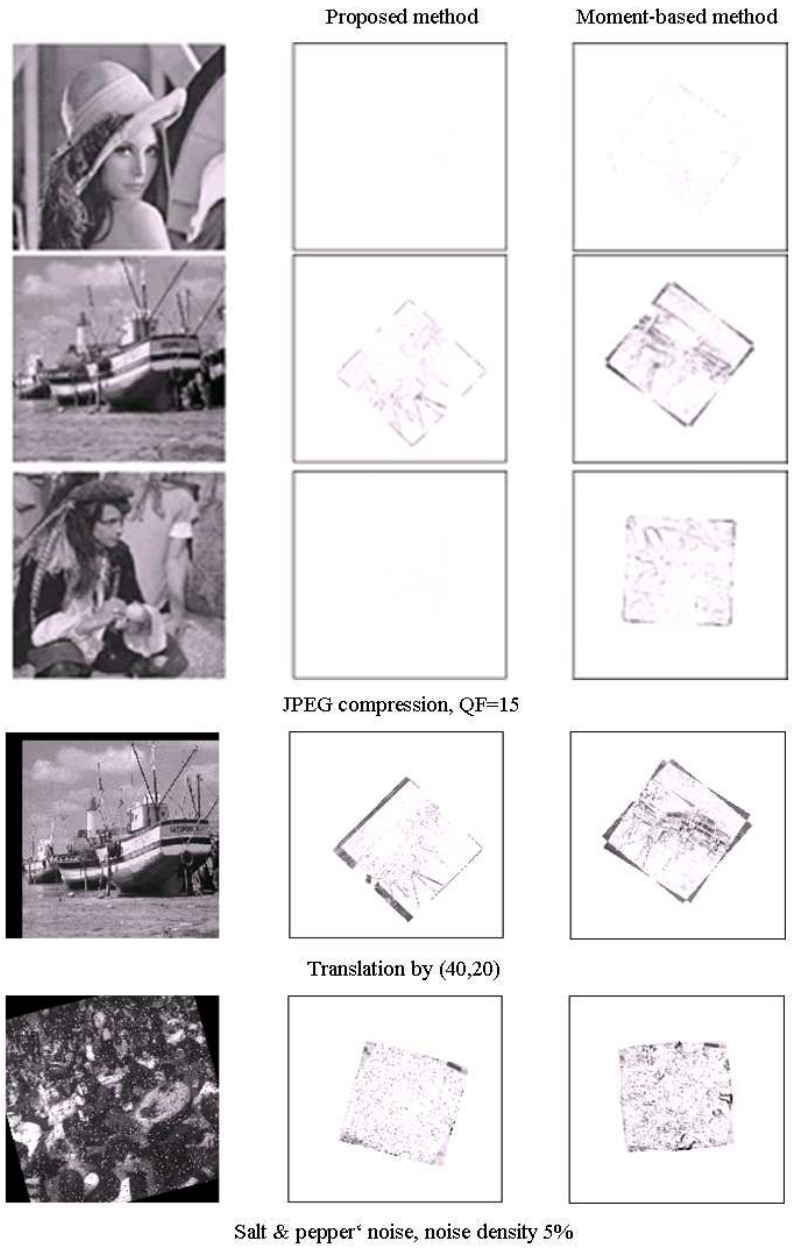



Figure 5.7: Evaluating the obtained normalization results: the obtained difference-images using the proposed technique and using the moment-based method in [65] for different distortions/degradations (JPEG compression, translation and 'Salt & pepper' noise contamination).

It can be seen in Fig. 5.6 and 5.7 that the difference-images of the proposed technique contain very few details. This demonstrates the accuracy of normalization and the robustness against several image-processing operations. Further, It can be seen from the results that the difference-images of the proposed method contain very few details compared to the technique presented in [65] and, hence, leads to more accurate normalization.

The normalization errors for the whole set of images using both techniques are obtained and presented in terms of their mean and standard deviation. The obtained results are plotted in Fig. 5.8. In this figure, the left graph shows the mean of the normalization parameters errors while the right graph shows the standard deviation of these errors.

It can be seen in that figure that the proposed technique has less normalization errors than the one in [65] and this confirms the normalization accuracy of the proposed technique.

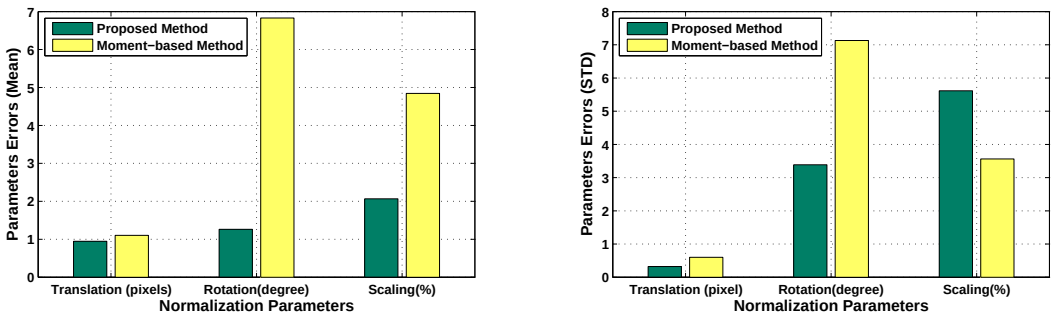



Figure 5.8: Evaluating the accuracy of the estimated normalization parameters: the left graph shows the means of the normalization parameters errors and the right graph shows the standard deviation of these errors.

## 5.4 An Image Watermarking Technique with Improved Resistance to Geometric Distortions using the Geometric Normalization Approach

As an example of the possible applications of the geometric normalization technique, presented in Section. 5.2, in this section, a blind watermarking technique for digital images is proposed. The main objective of the proposed technique is to ensure robustness to geometric distortions on top of the usual requirement of robustness to many common image-processing operations.

### 5.4.1 Introduction

Over the past years, digital watermarking techniques have been proposed as an effective method for copyright protection, as well as, for other applications. Such applications include embedding of watermarks that identify the rightful users of images, insertion of additional control or descriptive information, tampering detection (image authentication), controlling digital media access, or monitoring the data in broadcasting applications [67].

From the examination of any watermarking system, it can be noticed that it is very similar to a communication system. The goal of the watermarking system is basically the same as the goal of the communication system: to introduce some information into a medium, an image in our case, and then trying to extract it back. The watermark embedder could then be seen as the transmitter in the communication system, the watermark extractor as the receiver, and the host image as the communication channel [68].

In most applications, it is required that the watermark be invisible and robust. This means that the presence of the watermark should not interfere with the image

being watermarked and that it should be possible to detect/extract the embedded watermark after the watermarked image has undergone some processing operations. Any attempts to remove or destroy a watermark should result in severe degradation in fidelity before the watermark is lost. In particular, the watermark should be robust to most common image processing operations, e.g., compression, filtering, noise addition, etc. Moreover, a watermark system should be robust against geometric distortions, e.g., rotation, scale, translation, etc. Of course, there is a trade-off between watermark information capacity, watermark imperceptibility, and robustness [69], and this should be taken into consideration when designing a watermarking method.

So far, many watermarking algorithms have been proposed. Although most of these algorithms meet watermarking systems requirements, geometric distortions cause some difficulties to many of them, e.g., [70], [71], and [72]. Recently, the problem of robustness against geometric attacks, or in other words, the synchronization in watermark detection, has gained much importance. Resistance to geometric manipulations is important because such manipulations usually do not degrade the quality of the image while they are making the watermark undetectable. Geometric attacks affect only the synchronization between the embedded watermark and the watermark detector, whereas the watermark itself is usually still present in the data. One solution to this synchronization problem is to use exhaustive search to detect the watermark in blind watermarking systems [73]. However, this solution is not practical due to the computations required and the possibility of increase of false positive detection rate. Another solution in non-blind systems is based on using the original image in the watermark detection step. Unfortunately, this is impractical in cases where the original image is not available.

Several approaches have been proposed to deal with geometric attacks. There are mainly four different approaches proposed to deal with geometric attacks in the watermark detection: image normalization, invariant representation, synchronization

patterns and image features-based techniques. The basic idea of image normalization is to transform a given image into a standard form where this normalized form is independent of any possible geometric distortions applied on the image [65]. In the invariant representation techniques [74, 75] watermarks are embedded in invariant-transform domains in order to maintain invariance under rotation, scaling, and translation. Synchronization patterns techniques [76, 77], on the other hand, detect what the distortions were, and invert them before applying the watermark detector. In image features-based techniques [56, 78–80] significant features in images are used to help coping with geometric distortions by acting as reference points for watermark embedding/detection. Such features can be edges, corners, or any area with specific characteristics.

In the rest of this chapter, a watermarking technique, which can be used to increase the robustness of a watermarked image to geometric attacks, will be presented. This technique is based on the normalization technique presented in the first part of this chapter.

### 5.4.2 The Watermark Embedding/Detection Algorithm

The proposed technique is based on normalizing the image to be watermarked during the embedding process. The normalization parameters obtained during the embedding process and the watermark secret-key are used in watermark detection. For blind watermark detection, the same feature points are extracted from the watermarked image and used to transform the watermark into the normalized form before correlating it with the watermarked image. The basic idea behind the proposed technique is that if the watermarked image has been subjected to geometric distortions, then this will not affect the location of the feature points and thus it will be possible to obtain the normalized form of the watermark even though the

normalization parameter will be different in the embedding and extraction processes.

As in most watermarking systems, the watermarking technique considered here can be summarized in three main processes: watermark formatting, watermark embedding and watermark detection [80].

Let the host image be denoted by $I$, $W$ be the watermark to be embedded in the host image and $I^{'}$ the resulting watermarked image. The three processes of the proposed watermarking algorithm are explained as follows:

### Watermark Formatting

The watermark consists of a zero mean and unit variance two-dimensional pseudorandom sequence of 1s and $-1$s, which will be superimposed to the image's pixels, in the spatial domain. A key $K$ can be used as a seed for the random numbers generator. Each value of the sequence is spread in blocks of size $B \times B$ and all entries of each block have the same value (1 or $-1$).

### Watermark Embedding

The embedding process consists of the following two steps. In the first step, the normalization parameters are obtained for the image $I$ while in the second step, the generated watermark $W$ is embedded into the image. The description of each step is as follows.

1. *Normalization Parameter Estimation*:

   The normalization parameters $\mathbf{N_{p0}} = [\theta_0 \quad s_0 \quad t_{x_1 0} \quad t_{x_2 0}]$ are obtained using the image geometric normalization technique, presented in Section. 5.2, for the image $I$.

2. *Watermark Embedding*:

   The watermark $W$ is additively embedded in the luminance component of the

image. Each pixel of the image is modulated by the corresponding value of the embedding watermark multiplied by a watermark imperceptibility factor. A watermarked pixel is given by

$$I^{'}(x_1, x_2) = I(x_1, x_2) + \alpha(x_1, x_2).W(x_1, x_2), \tag{5.11}$$

where $\alpha(x_1, x_2)$ is a watermark imperceptibility factor at the image coordinates $(x_1, x_2)$, (see [80] for example) to control the visual image degradation due to watermark embedding.

**Watermark Detection**

In the detection process, the watermarked $I^{'}$ image which may have been subjected to geometric transformations and/or other image processing operations is available and the presence of the watermark $W$ has to be evaluated. The detection process, similarly to the embedding process, consists of the following steps:

1. *Watermark Transformation* As in watermark embedding process, normalization parameters $\mathbf{N_p} = [\theta \quad s \quad t_{x_1} \quad t_{x_2}]$ are obtained for the image $I^{'}$ and the obtained normalization parameters are used to transform the watermark $W$ to the normalized watermark $W_t$ as follows.

$$W_t(x, y) = T(t_{x_1} - t_{x_10}, t_{x_2} - t_{x_20}) + \frac{s}{s_0} R(\theta - \theta_0)W(x_1, x_2), \tag{5.12}$$

where $W(x_1, x_2)$ is a point of the watermark $W$, which is mapped to a point $W_t(x_1, x_2)$ in the transformed watermark $W_t$.

2. *Watermark Detection* The correlation value between the transformed test watermark $W_t$ and the possibly distorted image is calculated as

$$C = \frac{1}{MN} \sum_{x_1=1}^{M} \sum_{x_2=1}^{N} (I^{'}(x_1, x_2) - \overline{I_n^{'}})W_t^*(x_1, x_2) \tag{5.13}$$

where $M$ and $N$ are the image dimensions and $\overline{I'_n}$ is the local mean of a neighborhood of size $L * L$ with a center $(x_1, x_2)$. Then, the correlation value is compared with a threshold $T_d$ to decide whether the watermark is detected in the image or not. The threshold value is determined to maintain a certain watermark false positive probability, as in [80, 81].

The watermark embedding and detection processes are shown schematically in Fig. 5.9.

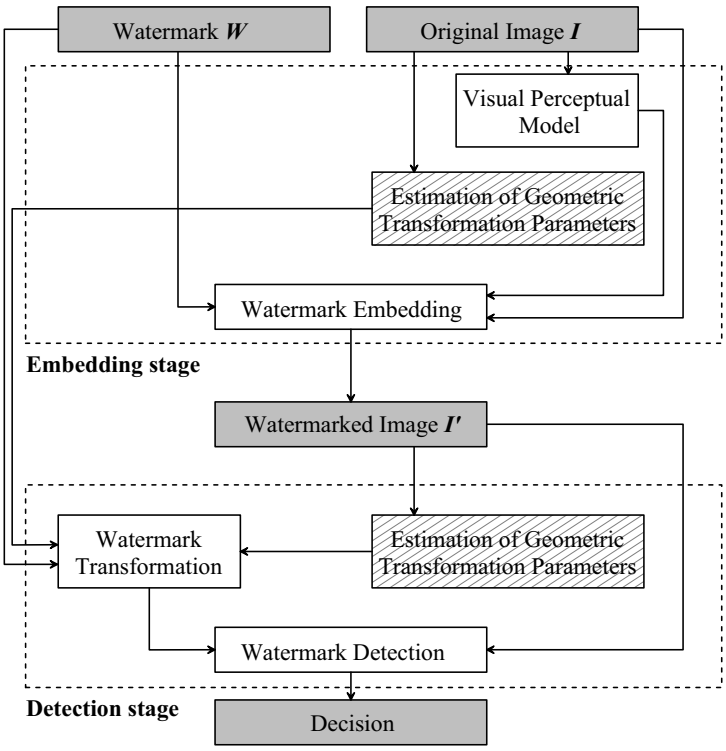



Figure 5.9: Watermark embedding and detection processes.

### 5.4.3 Performance Evaluation

In this section, the performance of the proposed technique is evaluated for different types of distortions. The 8-bit gray level 'Lena' and 'Bridge' images of size 512 ×

512 have been watermarked, geometrically distorted and noise has been added or some standard image processing operations, such as, cropping, brightness change, image compression, have been applied. Details of the applied distortions are listed in Table 5.1. The proposed technique has been used with a watermark block size

Table 5.1: Examples of distortions applied on the watermarked image.

| | |
|---|---|
| (a) | Scaling down by 70% and Rotation by $30^o$. |
| (b) | Averaging filter of size $3 \times 3$ and Rotation by $-5^o$. |
| (c) | Averaging filter of size $7 \times 7$ and Rotation by $-5^o$. |
| (d) | Brightness change by $-25$ and Rotation by $15^o$. |
| (e) | JPEG compression by Quality Factor of 35 and Rotation by $15^o$. |
| (f) | Cropping of 10 rows from one side (translation in $y$ direction by 5). |
| (g) | Cropping of 10 rows from one side (translation in $y$ direction by 5) and Rotation by $10^o$. |
| (h) | Cropping of 10 columns from one side (translation in $x$ direction by 5). |
| (i) | Cropping of 8 rows and columns from one side (translation in $x$ and $y$ directions by 4) and Rotation by $10^o$. |
| (j) | Rotation by $2^o$ and cropping of 7 rows/columns from each side. |
| (k) | Rotation by $0.5^o$. |

$B = 4$, the neighborhood's size $L = 5$ and the detection threshold $T_d = 0.2$. In order to evaluate visual distortions introduced by the embedding process, the PSNR of the difference between the original and watermarked images is used. The obtained PSNR when embedding a watermark into the images 'Lena' and 'Bridge' are 41.95 dB and 41.91 dB, respectively, which indicates that the image distortion introduced by watermark embedding is small. This can be also seen in Fig. 5.10, which shows the original and the watermarked images.

In order to evaluate the performance of the proposed technique, the following is considered:

Figure 5.10: Original (upper row) and watermarked (lower row) images('Lena' and 'Bridge').

- The difference between the actual geometric transformation applied on the image (listed in Table 5.1) and the transformation parameters obtained using the normalization technique (listed in the second column of Table 5.2).
- The obtained correlation values for detection, which are given in Table 5.2.

It can be noticed from the obtained results in Table 5.2 that despite the several types of distortions applied on the watermarked image, good estimation of the geometric transformations applied to the watermarked image can still be obtained as it can be seen from the first column in Table 5.2. Consequently, the obtained correlation values for detection are close to the ones obtained by using the exact (actual) transformation parameters, applied in the tests (listed in Table 5.1). Clearly, increased accuracy in the estimation of geometric distortion parameters, leads to higher correlation values.

A visual example showing the effect of applying geometric attacks on two images and the performance of the proposed technique can be seen in Fig. 5.11. The original and the watermarked images are shown in Fig. 5.11-a and b, respectively. It can be seen that the watermark embedding has not degraded the images (The obtained PSNR when embedding the watermarks into the images is about 41.9 dB). Fig. 5.11-c shows a simple geometric attacks of rotation by $2^o$ and cropping of 7 rows/columns from each side. There is no visible difference between the original and attacked images. However, as shown in Table 5.2, the watermark cannot be detected due to this slight rotation without using the normalization process. On the other hand, by using the proposed normalization technique, the correlation values are high enough to indicate the presence of the watermark. Fig. 5.11-d shows a simple geometric attacks of rotation by $0.5^o$. Since this is a small rotation, the watermark can be detected with and without the normalization process, as can be seen in Table 5.2. However, the correlation values when employing the normalization technique are higher.

Table 5.2: Effects of image distortions on the precision of the estimated transformation parameters and detection correlation values for 'Lena' and 'Bridge' images.

| Distortion (Table 5.1) | Estimated Angle during Detection | Scale | $t_{x_1}$ | $t_{x_2}$ | Correlation Value, using Eqn. 5.13: Proposed Method | Using Exact Trans. Parameters | No Geometric Adjustment |
|---|---|---|---|---|---|---|---|
| 'Lena' image | | | | | | | |
| (a) | $30.14^o$ | 0.698 | 0 | 1 | 0.5375 | 0.7540 | 0.0330 |
| (b) | $-4.96^o$ | 0.998 | 0 | 0 | 0.4727 | 0.4795 | 0.0525 |
| (c) | $-4.82^o$ | 1.000 | 0 | 0 | 0.1400 | 0.1452 | 0.0221 |
| (d) | $15.13^o$ | 0.996 | 0 | 0 | 0.4875 | 0.7641 | 0.0212 |
| (e) | $15.13^o$ | 0.996 | 0 | 0 | 0.4465 | 0.6397 | 0.0315 |
| (f) | $0.00^o$ | 1.000 | 0 | 5 | 1.0626 | 1.0626 | −0.0128 |
| (g) | $10.30^o$ | 0.998 | 0 | 5 | 0.4081 | 0.5155 | 0.0065 |
| (h) | $0.00^o$ | 1.000 | 5 | 0 | 1.0561 | 1.0561 | 0.0285 |
| (i) | $10.18^o$ | 0.999 | 4 | 4 | 0.4302 | 0.4645 | 0.0519 |
| (j) | $2.23^o$ | 1.000 | 0 | 0 | 0.6204 | 0.8452 | 0.0240 |
| (k) | $0.71^o$ | 0.997 | 0 | 0 | 0.5007 | 0.8658 | 0.3254 |
| 'Bridge' image | | | | | | | |
| (a) | $29.61^o$ | 0.699 | 0 | 1 | 0.5499 | 0.8672 | 0.0223 |
| (b) | $-5.06^o$ | 1.002 | 0 | 0 | 0.5199 | 0.5412 | −0.0015 |
| (c) | $-5.06^o$ | 1.002 | 0 | 0 | 0.1663 | 0.1698 | −0.0014 |
| (d) | $14.95^o$ | 0.998 | −1 | 0 | 0.5041 | 0.7494 | −0.0297 |
| (e) | $14.95^o$ | 0.998 | −1 | 0 | 0.5361 | 0.7426 | −0.0127 |
| (f) | $0.00^o$ | 1.000 | 0 | 5 | 1.1130 | 1.1130 | −0.0909 |
| (g) | $9.59^o$ | 0.999 | −1 | 5 | 0.2645 | 0.5798 | −0.0471 |
| (h) | $0.17^o$ | 1.000 | 4 | 0 | 0.5659 | 1.1207 | 0.0477 |
| (i) | $10.23^o$ | 0.998 | 3 | 4 | 0.2557 | 0.5059 | 0.0471 |
| (j) | $1.76^o$ | 0.998 | −1 | 0 | 0.5408 | 0.8892 | 0.1163 |
| (k) | $0.49^o$ | 0.998 | −1 | 0 | 0.5851 | 0.9053 | 0.3521 |

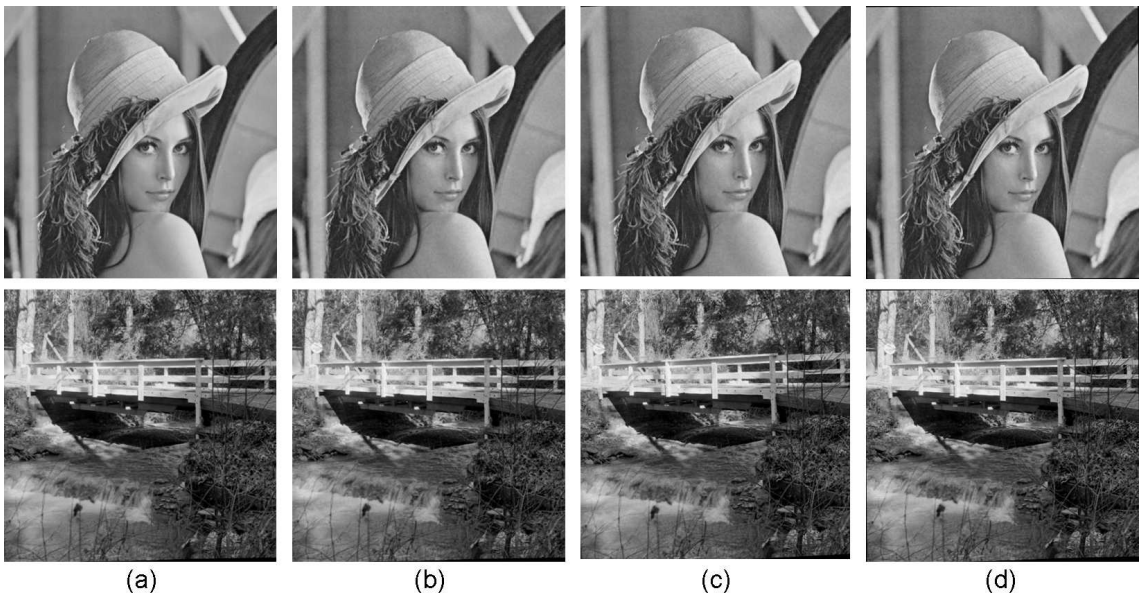



Figure 5.11: Effect of geometric attacks on the visual quality of 'Lena' (upper row) and 'Bridge' (lower row) images: (a) original image, (b) watermarked image, (c) the watermarked image has undergone rotation by $2^o$ and cropping of 7 pixels from each side, and (d) the watermarked image has undergone rotation by $0.5^o$.

## 5.5 Conclusions

In this chapter, a technique for image geometric normalization that relies on feature points of the image to perform the normalization process has been presented. The locations of a set of feature points, extracted from the image in hand, are used to estimate the normalization parameters needed to normalize the image. The performance of the proposed normalization technique was evaluated using various images with several types of degradations. Experimental results indicate that the proposed normalization technique gives good normalization accuracy and is robust against many image degradations such as image compression, brightness change, noise contamination and cropping.

A technique for image watermarking with robustness to geometric distortions has been proposed. The proposed technique is based on normalizing the image to be watermarked during watermark embedding and detection processes. Experimental analysis shows that the proposed watermarking technique achieves good robustness to such geometric distortions that include image translation, rotation, and scaling.

# Chapter 6

# Conclusions and Future Work

## 6.1 Conclusions

The focus of this dissertation has been the development of feature-based techniques for image registration and image geometric normalization. In Chapter 3, a feature point extractor that is based on scale-interaction of Mexican-hat wavelets and has the ability to handle images with scale change has been introduced. In Chapter 4, an image registration algorithm for registering geometrically distorted images, which may have partial overlap and, in addition, may have undergone additional degradations has been proposed. In Chapter 5, a technique for image geometric normalization based on the geometric properties of image feature points has been proposed. This technique has been used to enhance the robustness of an image watermarking technique with respect to geometric attacks.

### 6.1.1 Feature Point Extraction for Images with Scale Change

In Chapter 3, a feature point extractor that has the ability to handle images with scale change was introduced. The feature point extractor is based on scale-interaction of Mexican-hat wavelets. In order to adapt the feature point extraction process to

images scale change, a range of scaling factors is used for the Mexican-hat wavelet and feature points for different scaling factors are extracted.

The presented feature points extractor was tested using various images of different characteristics and its performance was evaluated using visual inspection (subjective test), repeatability rate and localization accuracy (objective test). It has been shown that the feature point extractor has reasonable invariance to image rotation and translation. Further, it can efficiently handle images with scale change. Experimental tests have been conducted to demonstrate its robustness and the obtained experimental results showed that the extracted feature points are robust against image degradations such as lossy image compression, image blurring and noise contamination.

### 6.1.2 Robust Image Registration

In Chapter 4, an algorithm for registration of geometrically distorted images, which may have partial overlap and may have undergone additional degradations has been proposed. The geometric distortions considered in the registration process are the global $2D$ affine transformations. The proposed algorithm consists of three main steps: extracting feature point using a feature point extractor based on scale-interaction of Mexican-hat wavelets, obtaining the correspondence between the feature points of the reference and the target images based on Zernike moments of neighborhoods centered on the feature points, and estimating the transformation parameters between the first and the second images using an iterative weighted least squares algorithm.

The performance of the proposed algorithm was evaluated using several types of images and degradations. It has been shown that the proposed algorithm gives excellent registration accuracy using several types of images, even in cases with partial

overlap between images. Further, experimental results indicate that the proposed algorithm can handle images of different scales effectively and is robust against many image degradations such as blurring, noise contamination, brightness change, etc.

### 6.1.3 Feature-based Image Normalization

In Chapter 5, a technique for image geometric normalization that relies on feature points has been proposed. The locations of a set of feature points, extracted from the image, are used to obtain the normalization parameters needed to normalize the image. The performance of the proposed technique was evaluated using several types of images and degradations. It was shown, based on the obtained experimental results, that the proposed technique yields good normalization accuracy. In addition, the robustness of the technique was examined and the obtained experimental results indicated that the proposed normalization is robust against many image degradations such as image compression, brightness change, noise contamination and image cropping.

As an example of the possible applications of the presented geometric normalization technique, a blind watermarking technique for digital images has been proposed. The main objective of the proposed technique is to ensure robustness to geometric distortions on top of the usual requirement of robustness to many common image-processing operations. In order to ensure robustness to geometric distortions, the proposed watermarking technique utilizes the geometric normalization technique to normalize the image, to be watermarked, during the embedding process. The normalization parameters obtained during the embedding process are used in watermark detection. For blind watermark detection, normalization parameters are obtained for the image and the obtained normalization parameters are used to transform the watermark into the watermark normalized form, which is correlated

with the possibly distorted watermarked image. Experimental analysis shows that the proposed watermarking technique achieves good robustness to geometric distortions that include image translation, rotation, and scaling.

## 6.2 Future Work

In what follows, some research topics that would extend the work presented in this dissertation are proposed.

### 6.2.1 Panoramic Image Stitching

In panoramic image stitching applications, two (or more) images that capture different views of the same scene, with an overlap region among the images, are combined together so that the stitched image contains no obstructive boundaries in the transition region while preserving the general appearance of the original images [82]. Clearly, a simple superimposing of one image on the other image produces visible artificial edges in the regions between the images, due to differences in camera [83]. The proposed image registration algorithm, in Chapter 4, can handle images that have partial overlap and it yields excellent registration accuracy using several types of images, including digital photography. In addition, it can effectively handle images of different scales (like zooming effect in cameras). Therefore, it can be well utilized in the applications of panoramic image stitching in the image registration stage. In the second stage, an image blending method for smoothing the transition between the images over their overlapping area is used. Many image blending method have been studied in literature and more work can still be done in this area.

### 6.2.2 Efficient Implementation of The Image Registration Algorithm

Nowadays, it is common to use high resolution images (i.e. their dimensions are in the order of thousand square pixels) in a wide range of applications. Registration of such high resolution images requires more computational resources to perform the registration algorithm stages. There are several opportunities for speeding up the implementation of the proposed algorithm by employing parallel processing or hardware implementations for computationally intensive tasks. Examples of the operations that can be accelerated include the following:

- The inherent parallelism of the filtering operation, performed during feature extraction process. This property can be beneficial by employing parallel processing, in hardware or software implementations where applicable.

- Efficient computation of Zernike moments. As the area of image increases, more computations are required for obtaining the Zernike moments of the feature point-centered neighborhoods. Therefore, the algorithm can be further accelerated by speeding up the computation of the Zernike moments. A good way is to employ hardware architecture to calculate Zernike moments as proposed in [84].

### 6.2.3 Possible Extensions of The Normalization Technique

The performance evaluation of the proposed image geometric normalization technique, in Chapter 5, shows that the proposed technique demonstrates good results in terms of normalization accuracy and robustness against many types of image degradations combined with geometric distortions. However, in order to meet other possible application requirements, some extensions of the proposed normalization technique can be investigated, for example:

- Relying on more feature points and employing a more sophisticated relationship between the points so that the geometric normalization technique can handle images that have undergone more complex geometric distortions, for example, affine transformation.

- Assigning additional weights to the feature points to control their contribution on the estimation of the normalization parameters. These weights could be assigned according to their detection response values, which determine their significance.

# Bibliography

[1] S. C. Pei and C. N. Lin, "Image normalization for pattern recognition," *Image and Vision Computing*, vol. 13, no. 10, pp. 711–723, 1995.

[2] Y. S. Avrithis, Y. Xirouhakis, and S. D. Kollias, "Affine-invariant curve normalization for object shape representation, classification and retrieval," *Machine Vision and Applications, Springer Verlag*, vol. 13, no. 2, pp. 80–94, 2001.

[3] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys (CSUR)*, vol. 24, no. 4, pp. 325–376, 1992.

[4] D. Shen and H. H. S. Ip, "Generalized affine invariant image normalization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 431–440, 1997.

[5] M. Sonka and J. M. Fitzpatrick, *Handbook of Medical Imaging: Volume 2, Medical Image Processing and Analysis*. SPIE Press, 2000.

[6] D. Salomon, *Transformations And Projections in Computer Graphics*. Springer-Verlag New York Inc, 2006.

[7] P. S. Heckbert, "Fundamentals of texture mapping and image warping," M.Sc. Thesis, University of California, Berkeley, 1989. [Online]. Available: http://citeseer.ist.psu.edu/150876. Last accessed: June 2007.

[8] B. Macq and J. D. E. J. Delp, "Benchmarking of image watermarking algorithms for digital rights management," *Proceedings of the IEEE*, vol. 92, no. 6, 2004.

[9] R. Haralick and L. Shapiro, *Computer and Robot Vision II.* Addison-Wesley, 1993.

[10] W. Forstner and E. Gulch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *Proceedings of ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland, June 1987, pp. 281–305.

[11] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of 4th Alvey Vision Conference*, 1988, pp. 147–151.

[12] S. Smith and J. Brady, "SUSAN - a new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, May 1997.

[13] B. Zitová, J. Kautsky, G. Peters, and J. Flusser, "Robust detection of significant points in multiframe images," *Pattern Recognition Letters*, vol. 20, no. 2, pp. 199–206, 1999.

[14] C. S. Kenney, B. S. Manjunath, M. Zuliani, M. G. A. Hewer, and A. V. Nevel, "A condition number for point matching with application to registration and postregistration error estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1437–1454, November 2003.

[15] L. M. G. Fonseca, "Automatic registration of remote sensing images based on multiple resolutions," Ph.D. dissertation, Computer Science , Instituto Nacional de Pesquisas Espaciais, SP, Brazil, 1999.

[16] H. Li, B. S. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Transactions on Image Processing*, vol. 4, no. 3, pp. 320–334, 1995.

[17] X. Dai and S. Khorram, "A feature-based image registration algorithm using improved chain-code representation combined with invariant moments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2351–2362, 1999.

[18] T. Lindeberg, "Scale-space theory: A basic tool for analysing structures at different scales," *Journal of Applied Statistics*, vol. 21, no. 2, pp. 224–270, 1994.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[20] K. Mikolajczyk, T. Tuytelaars, C. Schmid, and A. Z. et. al, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, 2005.

[21] B. Zitová and J. Flusser, "Image registration methods: a survey," *Proceedings of Image Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[22] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998. [Online]. Available: http://www.oup.co.uk/mediaj/hdb/Volume02/Issue01/. Last accessed: June 2007.

[23] J. P. W. Pluim, J. B. A. Maintz, and M. A. . Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, August 2003.

[24] H.-M. Chen, P. K. Varshney, and M. K. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 11, pp. 2445–2454, November 2003.

[25] B. S. Reddy and B. N. Chatterji, "An fft-based technique for translation, rotation, and scale-invariant image registration," *IEEE Transactions on Image Processing*, vol. 5, no. 8, pp. 1266–1271, August 1996.

[26] G. Wolberg and S. Zokai, "Robust image registration using log-polar transform," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Vancouver, BC, Canada, September 2000.

[27] H. Liu, B. Guo, and Z. Feng, "Pseudo-log-polar fourier transform for image registration," *IEEE Signal Processing Letters*, vol. 13, no. 1, pp. 17–20, 2006.

[28] M. Xia and B. Liu, "Image registration by "super-curves"," *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 720–732, May 2004.

[29] M. S. Yasein and P. Agathoklis, "Automatic and robust image registration using feature points extraction and zernike moments invariants," in *Proceedings of the Fifth IEEE International Symposium on Signal Processing and Information Technology*, Athens, Greece, December 2005, pp. 566–571.

[30] J. Flusser and B. Zitová, "Combined invariants to linear filtering and rotation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 13, pp. 1123–1136, 1999.

[31] J. Flusser and B. Zitová, "Combined invariants to convolution and rotation and their application to image registration," in *Proceedings of the Second International Conference on Advances in Pattern Recognition (ICAPR '01)*, London, UK, 2001, pp. 359–368.

[32] T. Kim and Y.-J. Im, "Automatic satellite image registration by combination of matching and random sample consensus," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 5, pp. 1111–1117, 2003.

[33] D. Fedorov, L. M. G. Fonseca, C. Kenney, and B. S. Manjunath, "Automatic registration and mosaicking system for remotely sensed imagery," in *Proceedings of SPIE 9th International Symposium on Remote Sensing*, Crete, Greece, September 2002, pp. 444–451.

[34] L. M. G. Fonseca and M. H. M. Costai, "Automatic registration of satellite images," in *Proceedings of Brazilian Symposium on Computer Graphics and Image Processing*, 1997, pp. 219–226.

[35] A. Goshtasby, "Template matching in rotated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 3, pp. 338–344, May 1985.

[36] P. Perez and N. Garcia, "Robust and accurate registration of images with unknown relative orientation and exposure," in *Proceedings of the IEEE International Conference on Image Processing*, 2005, pp. 1104–1107.

[37] F. Badra, A. Qumsieh, and G. Dudek, "Rotation and zooming in image mosaicing," in *Proceedings of the Fourth IEEE Workshop on Applications of Computer Vision WACV '98*, Princeton, NJ, USA, October 1998, pp. 50–55.

[38] C.-J. Liao and S.-Y. Chen, "Complementary retrieval for distorted images," *Pattern Recognition*, vol. 35, no. 8, pp. 1705–1722, 2002.

[39] G. D. Finlayson, B. Schiele, and J. L. Crowley, "Comprehensive colour image normalization," *Computer Vision ECCV'98*, pp. 475–490, 1998. [Online]. Available: http://www.springerlink.com/content/nhcjvku49b4jnn5r. Last accessed: June 2007.

[40] M.-K. Hu, "Visual pattern recognition by moment invariants," *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[41] J. Wood, "Invariant pattern recognition: a review," *Pattern Recognition*, vol. 29, no. 1, pp. 1–17, 1996.

[42] Y. Abu-Mostafa and D. Psaltis, "Image normalization by complex moments," *IRE Transaction on Pattern Analysis and Machine Intelligence*, vol. 7, no. 1, pp. 46–55, 1985.

[43] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496–513, 1988.

[44] A. Rosenfeld and A. C. Kak, *Digital Picture Processing, vol. 2.* New York: Academic Press, 1982.

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

[46] I. Rothe, H. Susse, and K. Voss, "The method of normalization to determine invariants," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 4, pp. 366–376, 1996.

[47] C. Yuccer and K. Oflazer, "A rotation, scaling and translation invariant pattern classification system," *Pattern Recognition*, vol. 26, no. 5, pp. 687–710, 1993.

[48] J. Liu and T. Zhang, "Matching and normalization of affine deformed image from regular moments," *Pattern Recognition Letters*, vol. 25, no. 14, pp. 1619–1631, 2004.

[49] Y. Zhang, C. Wen, Y. Zhang, and Y. C. Soh, "On the choice of consistent canonical form during moment normalization," *Pattern Recognition Letters*, vol. 24, no. 16, pp. 3205–3215, 2003.

[50] M. Gruber and K. Hsu, "Moment-based image normalization with high noise tolerance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 136–139, 1997.

[51] Y. S. Avrithis, Y. Xirouhakis, and S. D. Kollias, "Affine-invariant curve normalization for shape-based retrieval," in *Proceedings of the IEEE International Conference on Pattern Recognition (ICPR)*, Barcelona, Spain, September 2000, pp. 5015–5018.

[52] J. Cortadellas, J. Amat, and F. de la Torre, "Robust normalization of silhouettes for recognition applications," *Pattern Recognition Letters*, vol. 25, no. 5, pp. 591–601, 2004.

[53] B. Manjunath, C. Shekhar, and R. Chellappa, "A new approach to image feature detection with applications," *Pattern Recognition*, vol. 29, no. 4, pp. 627–640, 1996.

[54] A. Dobbins, S. W. Zucker, and M. S. Cynader, "Endstopping and curvature," *Vision Research*, vol. 29, no. 10, pp. 1371–1387, 1989.

[55] S. Bhattacharjee and M. Kutter, "Compression tolerant image authentication," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Chicago, USA, October 1998, pp. 435–439.

[56] M. Kutter, S. K. Bhattacharjee, and T. Ebrahimi, "Toward second generation watermarking schemes," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Kobe, Japan, October 1999, pp. 320–323.

[57] M. S. Yasein and P. Agathoklis, "A robust, feature-based algorithm for aerial image registration," in *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE 2007)*, Vigo, Spain, June 2007, pp. 1731–1736.

[58] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings Royal Society of London*, vol. 207, no. 1167, pp. 187–217, 1980.

[59] P. Tissainayagam and D. Suter, "Assessing the performance of corner detectors for point feature tracking applications," *Image and Vision Computing*, vol. 22, no. 8, pp. 663–679, 2004.

[60] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 489–497, 1990.

[61] S. X. Liao and M. Pawlak, "On the accuracy of zernike moments for image analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1358–1364, 1998.

[62] P. Hew, "Geometric and zernike moments," Department of Mathematics, The University of Western Australia," Diary, October 1996. [Online]. Available: http://citeseer.ist.psu.edu/hew96geometric.html. Last accessed: June 2007.

[63] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust statistics: the approach based on influence functions.* Wiley, 1986.

[64] E. Malis and E. Marchand, "Experiments with robust estimation techniques in real-time robot vision," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, October 2006, pp. 223–228.

[65] M. Alghoniemy and A. H. Tewfik, "Geometric invariance in image watermarking," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 145–153, 2004.

[66] M. S. Yasein and P. Agathoklis, "A feature-based image normalization technique for handling geometric distortions," in *Proceedings of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, California, USA, October–November 2005, pp. 884–887.

[67] M. Swanson, M. Kobayashi, and A. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1064–1088, 1998.

[68] J. R. Martin and M. Kutter, "Information retrieval in digital watermarking," *IEEE Communications magazine*, vol. 39, no. 8, pp. 110–116, Aug 2001.

[69] C. I. Podilbuk and E. J. Delp, "Digital watermarking: Algorithms and applications," *IEEE Signal Processing Magazine*, vol. 18, no. 4, pp. 33–46, 2001.

[70] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[71] S. D. Lin and C.-F. Chen, "A robust DCT-based watermarking for copyright protection," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 3, pp. 415–421, August 2000.

[72] M. S. Yasein and P. Agathoklis, "A Wavelet-based blind and readable image watermarking algorithm," in *Proceedings of the 36th Asilomar Conference on*

*Signals, Systems and Computers*, California, USA, Nov 2002, pp. 1215–1219, vol. 2.

[73] J. Lichtenauer, I. Setyawan, T. Kalker, and R. Lagendijk, "Exhaustive geometrical search and the false positive watermark detection probability," in *Proceedings of SPIE Security and Watermarking of Multimedia Contents V*, vol. 5020, Santa Clara, CA, USA, June 2003, pp. 203–214.

[74] J. O. Ruanaidh and T. Pun, "Rotation scale and translation invariant spread spectrum digital image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 303—317, 1998.

[75] D. Zheng, J. Zhao, and A. E. Saddik, "RST-invariant digital image watermarking based on log-polar mapping and phase correlation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 753–765, 2003.

[76] S. Pereira and T. Pun, "Robust template matching for affine resistant image watermarks," *IEEE Transactions on Image Processing*, vol. 9, no. 6, pp. 1123–1129, 2000.

[77] X. G. Kang, J. W. Huang, Y. Q. Shi, and Y. Lin, "A DWT-DFT composite watermarking scheme robust to both affine transform and JPEG compression," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 776–786, 2003.

[78] P. Bas, J.-M. Chassery, and B. Macq, "Geometrically invariant watermarking using feature points," *IEEE Transactions On Image Processing*, vol. 11, no. 9, pp. 1014–1028, 2002.

[79] C.-W. Tang and H.-M. Hang, "A feature-based robust digital image watermarking scheme," *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 950–959, 2003.

[80] D. Simitopoulos, D. E. Koutsonanos, and M. G. Strintzis, "Robust image watermarking based on generalized radon transformations," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 732–745, 2003.

[81] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "Threshold selection for correlation-based watermark detection," in *Proceedings of COST 254 Workshop on Intelligent Communications*, L'Aquila, Italy, June 1998, pp. 67–72.

[82] M.-S. Su, W.-L. Hwang, and K.-Y. Cheng, "Analysis on multiresolution mosaic images," *IEEE Transactions on Image Processing*, vol. 13, no. 7, pp. 952–959, July 2004.

[83] A. Zomet, A. Levin, S. Peleg, and Y. Weiss, "Seamless image stitching by minimizing false edges," *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 969–977, April 2006.

[84] L. Kotoulas and I. Andreadis, "Real-time computation of Zernike moments," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 801–809, June 2005.

# Appendix A

# Derivation of the formula for the Mexican-hat wavelet in the spatial domain from its frequency domain form

In this appendix, we derive equation 3.1 for the Mexican-hat wavelet in the spatial domain ($Mex(\mathbf{x}, s_m)$), from its frequency domain form.

According to equation 3.5, the Mexican-hat wavelet can be defined in the frequency domain as

$$Mex_f(\mathbf{x_f}, s_m) = \sigma^3(x_{f_1}^2 + x_{f_2}^2)\, e^{\left(-\frac{1}{2}\sigma^2(x_{f_1}^2 + x_{f_2}^2)\right)}, \tag{A.1}$$

where $\mathbf{x_f} = (x_{f_1}, x_{f_2})$ is the 2D spatial-frequency variable.

Taking the inverse Fourier transform of $Mex_f(\mathbf{x_f}, s_m)$, the Mexican-hat wavelet in the spatial domain ($Mex(\mathbf{x}, s_m)$) can be obtained by

$$Mex(\mathbf{x}, s_m) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Mex_f(\mathbf{x_f}, s_m) e^{j(x_{f_1} x_1 + x_{f_2} x_2)} dx_{f_1} dx_{f_2} \tag{A.2}$$

Hence,

$$
\begin{aligned}
Mex(\mathbf{x}, s_m) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma^3 (x_{f_1}^2 + x_{f_2}^2)\, e^{\left(-\frac{1}{2}\sigma^2 (x_{f_1}^2 + x_{f_2}^2)\right)} e^{j(x_{f_1} x_1 + x_{f_2} x_2)} dx_{f_1} dx_{f_2} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma^3 x_{f_1}^2\, e^{-\frac{1}{2}\sigma^2 x_{f_1}^2}\, e^{j x_{f_1} x_1}\, e^{-\frac{1}{2}\sigma^2 x_{f_2}^2}\, e^{j x_{f_2} x_2} dx_{f_1} dx_{f_2} \\
&+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma^3 x_{f_2}^2\, e^{-\frac{1}{2}\sigma^2 x_{f_2}^2}\, e^{j x_{f_2} x_2}\, e^{-\frac{1}{2}\sigma^2 x_{f_1}^2}\, e^{j x_{f_1} x_1} dx_{f_1} dx_{f_2} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \sigma^3 x_{f_1}^2\, e^{-\frac{1}{2}\sigma^2 x_{f_1}^2}\, e^{j x_{f_1} x_1} dx_{f_1} \,.\, \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2 x_{f_2}^2}\, e^{j x_{f_2} x_2} dx_{f_2} \\
&+ \frac{1}{2\pi} \int_{-\infty}^{\infty} \sigma^3 x_{f_2}^2\, e^{-\frac{1}{2}\sigma^2 x_{f_2}^2}\, e^{j x_{f_2} x_2} dx_{f_2} \,.\, \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sigma^2 x_{f_1}^2}\, e^{j x_{f_1} x_1} dx_{f_1} \\
&= \frac{\sigma^3}{2\pi} \left( -\sqrt{2\pi} \frac{x_1^2}{\sigma^5}\, e^{-\frac{1}{2}\frac{x_1^2}{\sigma^2}} + \frac{\sqrt{2\pi}}{\sigma^3}\, e^{-\frac{1}{2}\frac{x_1^2}{\sigma^2}} \right) \,.\, \left( \frac{\sqrt{2\pi}}{\sigma}\, e^{-\frac{1}{2}\frac{x_2^2}{\sigma^2}} \right) \\
&+ \frac{\sigma^3}{2\pi} \left( -\sqrt{2\pi} \frac{x_2^2}{\sigma^5}\, e^{-\frac{1}{2}\frac{x_2^2}{\sigma^2}} + \frac{\sqrt{2\pi}}{\sigma^3}\, e^{-\frac{1}{2}\frac{x_2^2}{\sigma^2}} \right) \,.\, \left( \frac{\sqrt{2\pi}}{\sigma}\, e^{-\frac{1}{2}\frac{x_1^2}{\sigma^2}} \right) \\
&= \frac{\sigma^3}{2\pi} \left( \frac{-2\pi x_1^2}{\sigma^6}\, e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)} + \frac{2\pi}{\sigma^4}\, e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)} \right) \\
&+ \frac{\sigma^3}{2\pi} \left( \frac{-2\pi x_2^2}{\sigma^6}\, e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)} + \frac{2\pi}{\sigma^4}\, e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)} \right) \\
&= \frac{\sigma^3}{2\pi} \left( \frac{-2\pi x_1^2}{\sigma^6} - \frac{2\pi x_2^2}{\sigma^6} + \frac{2\pi}{\sigma^4} + \frac{2\pi}{\sigma^4} \right)\, e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)} \\
&= \frac{1}{\sigma} \left( 2 - \frac{x_1^2 + x_2^2}{\sigma^2} \right)\, e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)} \qquad \text{(A.3)}
\end{aligned}
$$

Therefore, in the spatial domain, the Mexican-hat wavelet can be expressed as

$$
Mex(\mathbf{x}, s_m) = \frac{1}{\sigma} \left( 2 - \frac{x_1^2 + x_2^2}{\sigma^2} \right) e^{\left(-\frac{1}{2}\frac{x_1^2 + x_2^2}{\sigma^2}\right)}, \qquad \text{(A.4)}
$$

# Appendix B

# Additional examples of image registration, for the technique presented in Chapter 4

In this appendix, more examples that demonstrate the performance of the registration algorithm, presented in Chapter 4, are shown. Various types of images with different characteristics are used.

Each example shows the two images with the paired feature points superimposed on each image, the registered image overlaid on the reference one, and the feature points of the registered (transformed) image superimposed on the feature points of the reference image.

The quality of registration are evaluated using the registration errors ($D_M$ and $D_{STD}$), which are be determined as in (4.22), and are indicated under each figure.

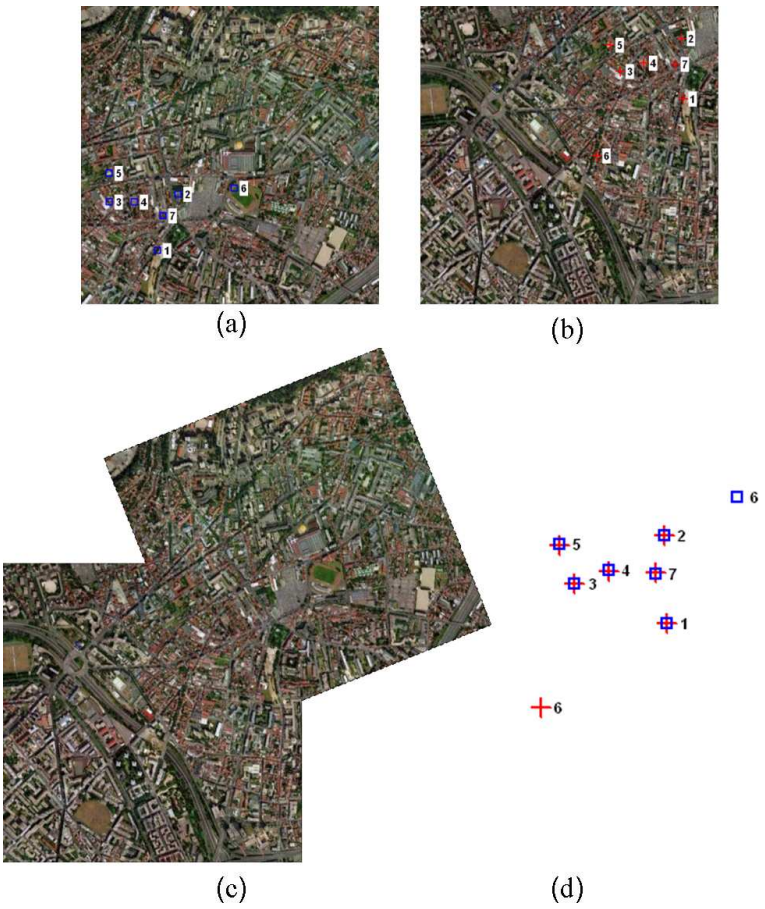



Figure B.1: Additional registration Example 1 (registration of two aerial images having a small overlap area (21% of the image area)), adapted from Google Earth. The obtained registration error ($D_M$ and $D_{STD}$) are 0.65 and 0.26 pixel, respectively: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

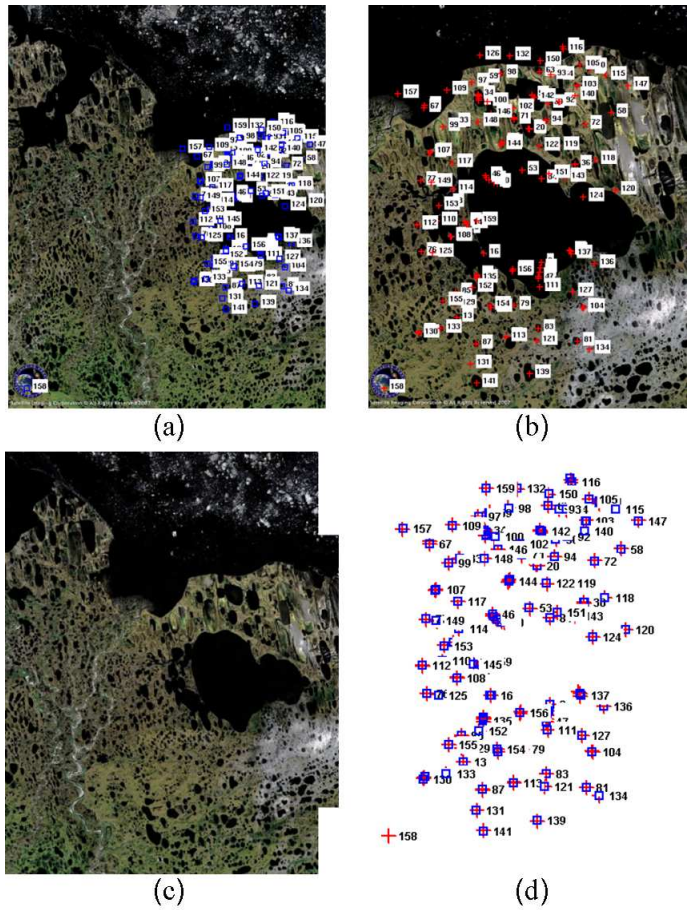



Figure B.2: Additional registration Example 2 (registration of two satellite images having different scale of Alaska region (Landsat)), available at http://www.satimagingcorp.com/gallery-landsat.html. The obtained registration error ($D_M$ and $D_{STD}$) are 0.94 and 0.66 pixel, respectively: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

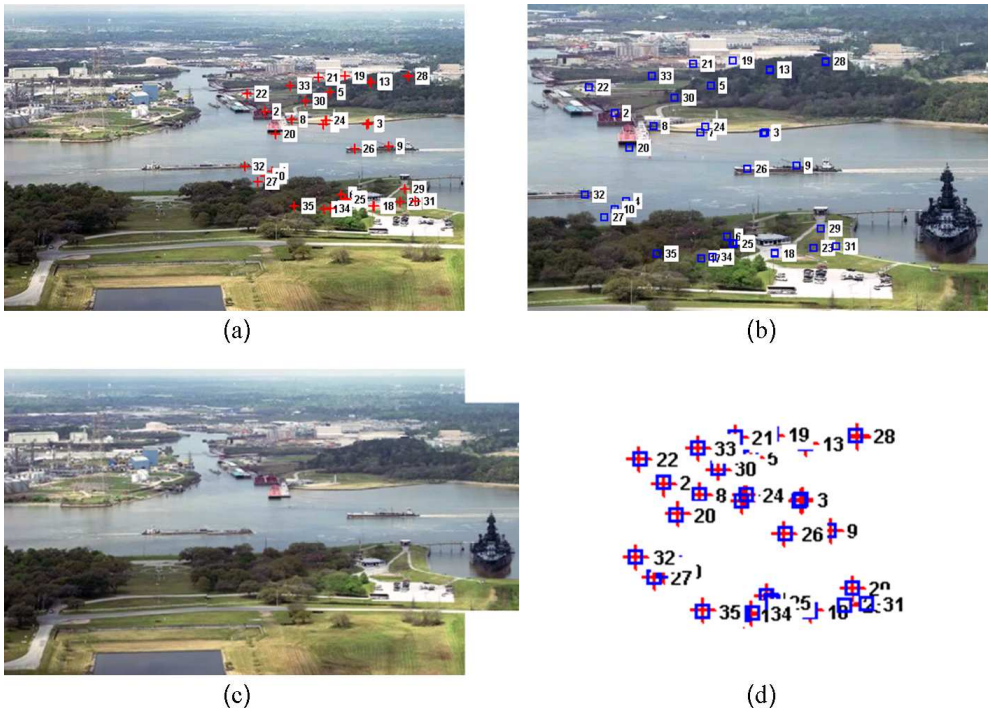



Figure B.3: Additional registration Example 3 (registration of two photographs having different scale of an outdoor scene), available at http://www.gigapxl.org/gallery-ViewSanJacinto.htm. The obtained registration error ($D_M$ and $D_{STD}$) are 0.30 and 0.12 pixel, respectively: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

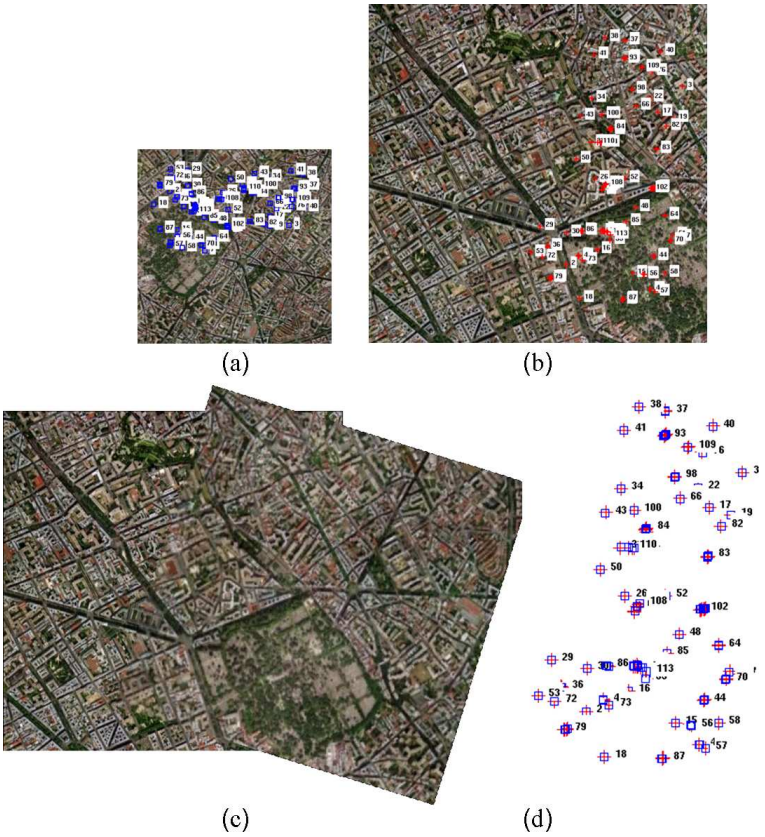



Figure B.4: Additional registration Example 4 (registration of two aerial images having different scale), adapted from Google Earth. The obtained registration error ($D_M$ and $D_{STD}$) are 0.26 and 0.11 pixel, respectively: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

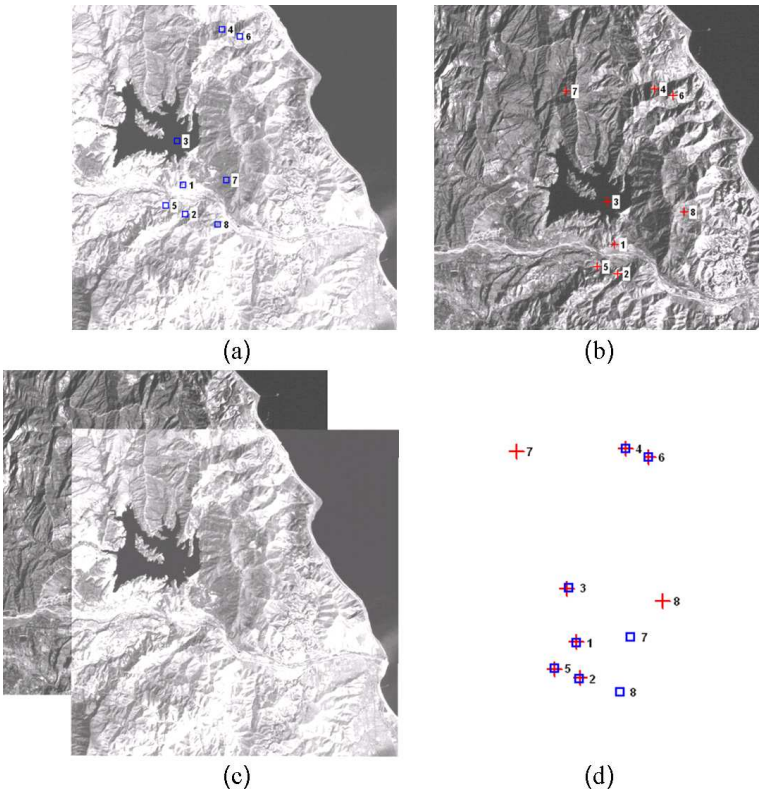



Figure B.5: Additional registration Example 5 (registration of two Landsat TM (thematic mapper) multi-spectral satellite images (Landsat TM Band 5 and 7)), available at http://vision.ece.ucsb.edu/registration/satellite/testimag/index.htm. The obtained registration error ($D_M$ and $D_{STD}$) are 1.67 and 0.81 pixel, respectively: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).

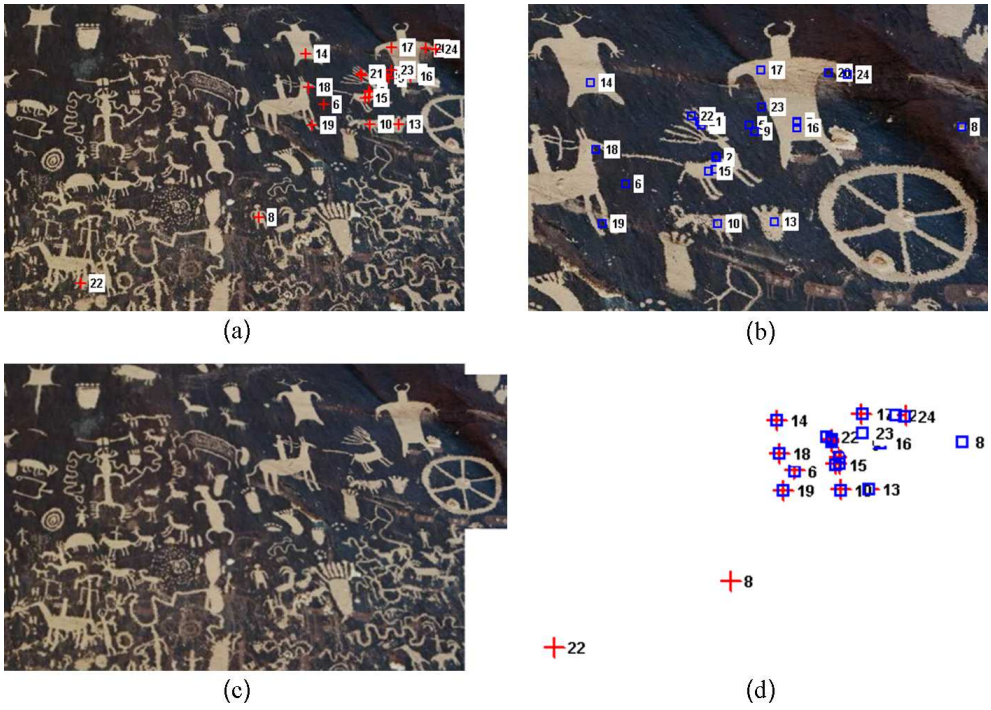



Figure B.6: Additional registration Example 4 (registration of two photographs having different scale of petroglyph panel), available at http://www.gigapxl.org/gallery-Newspaper.htm. The obtained registration error ($D_M$ and $D_{STD}$) are 0.63 and 0.84 pixel, respectively: (a) and (b) The two image with the paired feature points superimposed on each image; (c) The registered image overlaid on the reference image; (d) The feature points of the registered (transformed) image (represented by squares) superimposed on the feature points of the reference image (represented by crosses).