MONASH University

An Alignment-free Approach for Image Matching Applications

by

Komal Komal

Bachelor of Engineering in Electronics and Communications, DCRUST, Murthal, India

Master of Technology in Electronics and Communications, NIT, Kurukshetra, India

A thesis submitted to the

Faculty of Information Technology

for fulfilment of the requirement for the Degree of

Doctor of Philosophy (0190)

Monash University

Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Komal Komal

April, 2018

Acknowledgement

I would like to express immeasurable gratitude to the following who with the best of their guidance and support have contributed in making this study possible:

- I put on record the wisdom of my supervisors Dr Nandita Bhattacharjee, Dr David Albrecht and Prof Bala Srinivasan for being excellent supervisors, great mentors and for their invaluable patience, motivation and immense knowledge.

- I submit the role of the members of the committee for examination constituting Prof Frada Burstein, Dr Ron Steinfeld, Dr Maria Indrawan and Associate Prof Carsten Rudolph who provided their imperative skills and talents by way of correction and ideas shared which enabled me to finalize this work.

- I am grateful to the Faculty of Information Technology (FIT) and the Monash Institute of Graduate Research (MIGR) for providing me financial soundness in the form of a scholarship for my PhD.

- I am thankful to my friends Yathindu Hettiarachchige, Anchal Sareen, Yang (Kelvin) Li, Bo Chen, Mohammad Shamshur Rahman, Vidya Saikrishna and Jishan-e-Giti; and Faculty of information technology staff Sue Mckemmish, Danette Deriane and Helen Cridland who directly or indirectly extended their support that helped me a lot to complete my work in time. Also, I am thankful to Viranga Ratnaike for proofreading my thesis.

- Last but not least, I would like to thank my wonderful parents and family for their love and unconditional support that kept me going in extremely difficult and excruciating times. A special thank you to my partner Ankit Singh for his care, support, patience and prayers to complete this research work. Finally, my big hug to my nephew Remant and nieces Manyata and

Inayat who brighten up all my days. I am sure that I would not have been able to finish this study without their well wishes.

Komal Komal

Monash University

April, 2018

List of Publications

- **K. Komal, N. Bhattacharjee, D. Albrecht and B. Srinivasan**, *"Transformational approach for alignment-free image matching applications"*, in 15th International Conference on Advances in Mobile Computing and Multi Media, Austria, 2017, pp. 49-57.

- **K. Komal, D. Albrecht, N. Bhattacharjee and B. Srinivasan**, *"A region-based alignment-free partial fingerprint matching"*, in 14th International Conference on Advances in Mobile Computing and Multi Media, Singapore, 2016, pp. 63-70.

- **O. Zanganeh, K. Komal, N. Bhattacharjee, D. Albrecht and B. Srinivasan**, *"Quality Controlled Region-Based Partial Fingerprint Recognition"*, in Journal of Mobile Multimedia, Rinton press, 2018, Vol. 14-2, pp. 123-156.

Abstract

In computer vision and image processing, alignment of images plays an important role in different day to day life applications such as image registration and image matching applications. In image registration applications, alignment brings similar points of images into correspondence, while in image matching applications alignment acts as an intermediate step and it helps in assessing the similarity between two images. Medical image diagnosis and non-destructive testing for weld defect identification are two important examples of image registration applications. In such applications, images are aligned and then a manual inspection is done on aligned images to make decisions. On the other hand, some of the examples of applications which automatically match images are logo detection, object recognition, image retrieval, biometric authentication and identification amongst others. In these applications, images are first aligned and then matched automatically for recognition. The motivation of this thesis is derived from this increasing demand for accurate and efficient image registration and matching in different applications.

Conventionally, images are aligned by using the extracted features of images. However, the extraction of the features is a difficult and time consuming task as image processing is required for detection of features. In the case of distorted or poor quality images, the accurate detection of features can be difficult, whereas, in partial images, enough features required to confidently align images may be absent. Another method of aligning images is by using pixel based information of images. These methods perform well in the cases of partial and low quality images, however, are computationally exhaustive. The pixel based methods, using different transforms such as Radon or Fourier transforms, align images by computing the parameters between them. However, a pixel-to-pixel correspondence between images is mandatory for these methods. Such a correspondence may not exist due to many factors such as noise in images, available images being partial,

different regions of an image having different quality and the presence of other objects in the image. In such cases, the known methods of aligning images may fail. Therefore, to deal with these challenges, two tasks need to be performed regardless of whether images have pixel-to-pixel correspondence or not: (i) for image registration applications, compute the parameters between images not only accurately but also efficiently to align them; and (ii) for image matching applications where alignment acts as an intermediate step, accurately and efficiently match images without aligning them.

In the case, when pixel-to-pixel correspondence exists, the thesis proposes an efficient transform based method to compute the transformation parameters between two images for aligning them. The proposed approach is verified on test images of varying classes and results demonstrate that the proposed method can compute the parameters accurately and efficiently as compared to the existing approaches. In order to efficiently match images, the thesis investigates whether matching can be done without computing all the transform parameters or in other words without aligning them. This question is answered by mapping an image into a rotation invariant representation. However, the existing similarity metrics cannot be used to compute the similarity between images in this representation. A new similarity metric is proposed to compute the similarity between images. The effectiveness of the proposed approach is tested on logo datasets. Results indicate that the approach can achieve the same matching accuracy as efficiently as that of existing transform parameter based methods. Also, the approach is resilient even when images contain around 20% noise.

The proposed methods of computing the parameters and matching without alignment have a basic requirement that images should have pixel-to-pixel correspondence. These methods fail when partial or fragmented parts of full images are given as input for matching. However, in those images, there may be a few regions of images which have the required correspondence. Therefore, for efficient

aligning or matching of such images, we may need to extract those common regions efficiently. Since the images are unaligned, it is a time consuming process to extract the common region by using existing techniques. To extract common regions between images efficiently, we have proposed an alignment-free common region extraction approach. The extracted common region satisfies the required correspondence and can be used to compute the transformation parameters, if required, for aligning. The proposed approach can compute the parameters even when images do not have exact pixel-to-pixel correspondence, while most of the existing methods fail to perform. The extracted common regions can also be used for matching images which do not have exact pixel-to-pixel correspondence. For accurate and efficient image matching, a region-based alignment-free method is proposed. A single region between images might not be enough to accurately match them as a small part of given information is used by single region. Therefore, to improve accuracy, a multiple-region based method is proposed that can use most of the available information. The effectiveness of the proposed approach is studied by applying the approach to fingerprint identification systems as the acquisition of fingerprints can result in partial and low quality images. As compared to the state-of-the-art matching methods, the proposed approach not only improves the matching accuracy but also improves the computation time on the same test datasets.

In summary, the thesis has proposed a new paradigm by way of a transformational approach for image matching applications that can register and match images with improved accuracy and efficiency and without explicitly extracting image features. In addition, the proposed method is application independent and applicable to both full and partial images.

Contents

1 Introduction **1**

1.1 Preamble 1

1.2 Issues and Challenges 3

1.3 Motivation 8

1.4 Research Objectives 10

1.5 Research Contributions 11

1.6 Thesis Organization 13

2 Background and Related Work **15**

2.1 Preamble 15

2.2 Image Matching Systems 16

2.3 Categorization of Image Matching Techniques 18

2.4 Feature based Techniques 19

2.4.1 Feature Detection 20

2.4.2 Feature Matching 26

2.4.3 Computing Similarity 40

2.4.4 Summary of the Feature Based Techniques 41

2.5 Pixel based Techniques 42

2.5.1 Similarity Metrics 43

2.5.2 Search Methods 49

2.5.3 Types of Pixel Based Matching Techniques 50

2.5.4 Summary of Pixel Based Techniques 63
2.6 Hybrid Techniques 63
2.7 Learning Based Techniques 66
2.8 Comparison of Feature based and Pixel based Techniques 68
2.9 Conclusion 73

3 Recovery of Similarity Transformation Parameters 75
3.1 Preamble 75
3.2 Parameters to Compute 80
3.3 Transform for Computing Parameters 81
3.3.1 Rotation Property of Radon transform 83
3.4 Method of Computing Parameters 84
3.5 Computational Time 91
3.6 Experimental Results 93
3.7 Conclusion 102

4 Alignment-free Approach for Image Matching Applications 103
4.1 Preamble 103
4.2 Need for a Rotation Invariant Representation 106
4.3 Rotation Invariant Representation 108
4.4 Image Matching Algorithm 112
4.5 Computational Time 117
4.6 Performance Evaluation 118
4.6.1 Effect of the Number of Projections on Precision-Recall 123
4.7 Conclusion 125

5 Similarity Transformation Parameters Recovery in Partial Images 127

5.1 Preamble . . . 127
5.2 Partial Image Limitations . . . 129
5.3 Prerequisite for Computing Parameters . . . 132
5.4 Region as a Feature . . . 133
5.4.1 Shape of Region . . . 134
5.5 Method of Common Region Identification . . . 137
5.5.1 Optimisation of Common Region Extraction Method . . . 139
5.6 Method of Computing Parameters using a Common Region . . . 141
5.7 Conclusion . . . 144

6 Region-based Alignment-free Partial Image Matching 146
6.1 Preamble . . . 146
6.2 Issues in Partial Image Matching . . . 148
6.3 Partial Image Matching Requirements . . . 151
6.4 An Approach to Partial Image Matching . . . 152
6.4.1 Features in Partial Images . . . 153
6.4.2 Alignment-free Common Region Identification . . . 156
6.4.3 Similarity Computation between Unaligned Partial Images . 157
6.5 Performance of the Alignment-free Matching Approach . . . 162
6.5.1 Effect of the Number of Regions on Accuracy . . . 166
6.5.2 Effect of Size of Region on Accuracy . . . 168
6.5.3 Performance Evaluation . . . 169
6.5.4 Effect of Noise on Accuracy . . . 175
6.5.5 Effect of the Number of Radon Projections on Accuracy . . 176
6.5.6 Computational Complexity . . . 178
6.6 Conclusion . . . 179

7 Conclusion 181
7.1 Contributions of Thesis . . . 182

7.2 Future Work 184

List of Tables

2.1 Performance of the existing fingerprint matching techniques (in terms of EER) on the FVC2002DB1 dataset 39

2.2 Advantages and disadvantages of existing fingerprint matching techniques 69

3.1 The errors in parameter estimation for the Lena and Barbara images by using the proposed approach at different SNR with Gaussian noise 96

3.2 The estimated parameters for the Lena image by using the proposed approach as compared to the real parameters 96

3.3 The estimated parameters for the Barbara image by using the proposed approach as compared to the real parameters 97

3.4 The estimated parameters by using the proposed approach on the Barbara and Lena images at different densities of salt and pepper noise varying from 0 to 0.20 with an increment of 0.04 (Barbara images with varying densities is shown in Figure 3.8). 97

3.5 The errors in parameter estimation for the hand computed tomography image (shown in Figure 3.9) at different levels of both salt and pepper and Gaussian noise 99

3.6 The errors in parameters estimation for the fingerprint image (shown in Figure 3.10) at different levels of both salt and pepper and Gaussian noise 101

6.1 The accuracy achieved by the proposed method on different fingerprint datasets and for three different diameters (71,101,141 pixels) of a circular region by taking one of the impressions as sensed image. 170
6.2 The accuracy achieved for the FVC2002DB1-DB4 datasets by using the proposed method and by taking different impressions of each person as a sensed image 172
6.3 The accuracy achieved for the FVC2006DB2 dataset by using the proposed method and by taking different impressions of each fingerprint as sensed images 172
6.4 Comparison of the EER of the proposed method with the EER of existing methods on the FVC2002DB1 dataset 173
6.5 Comparison of the EER of the proposed approach with the EER of existing methods on the FVC2002DB2 dataset 173
6.6 Comparison of the EER of the proposed method with the EER of existing methods on the FVC2002DB3 dataset 174
6.7 Comparison of the EER of the proposed method with the EER of existing methods on the FVC2002DB4 dataset. 174
6.8 Comparison of the EER of the proposed approach with the EER of existing methods on FVC2006DB2 dataset. 174
6.9 The accuracy of the proposed approach at different densities of salt and pepper noise on the FVC2002DB1-DB4 and FVC2006DB2 datasets 176
6.10 The accuracy of the proposed approach with respect to the number of projections on the FVC2002DB1-DB4 and FVC2006DB2 datasets 177

List of Figures

1.1 The block diagram of research objectives 10

2.1 The block diagram of an image matching process 17

2.2 Corner detection in an image patch by using a twelve-point segment test [1] 22

2.3 In each octave the same image is convolved with a Gaussian of increasing scale, and finally these smoothed images are subtracted for producing a Difference-of-Gaussian [2]. 24

2.4 Ridge ending and bifurcation demonstrated in terms of binary pixels 25

2.5 The representation of a minutia (ridge ending and ridge bifurcation) [3] 26

2.6 An illustration of the K-plets in a fingerprint [4]. 29

2.7 A section of the cylinder around a minutia 'm'. All the minutiae involved in the construction of cylinder are shown [5]. 31

2.8 Sampling points arranged in a circular pattern around a minutia [6]. 33

2.9 A keypoint scale invariant feature transform (SIFT) descriptor [2]. . 33

2.10 BRISK descriptor [7] 35

2.11 Block diagram of the feature based transformation parameter computation approach [8]. 39

2.12 Pixel based matching methods: small windows of images are aligned by using normalised cross correlation (middle row) and phase correlation (last row). The global maximum identified in both graphs is the matching location. [9] 44
2.13 Block diagram of the image alignment method with or without using image features [10] 53
2.14 (a) Directional texture rotated at 45°; (b) Variance of Radon projections at different angles; (c) Second derivative of (b). The local maximum exists at 45° [11]. 59
2.15 The old and new photographs of a mosaic (top row) and mutual information criteria (bottom row). The mutual information gets maximised at point A instead of point B which is selected manually. The mistake is made due to the poor quality of the images [9]. . . . 62
2.16 System diagram of fingerprint authentication by fingercode [12] . . 65

3.1 The coloured block represents the thesis objective addressed in this chapter. 79
3.2 The computation of the Radon transform [13] 82
3.3 The 2π Radon transform rotation property: the rotation of an image by angle ϕ_0 is converted into the circular shift of ϕ_0 in the 2π Radon transform [14]. 84
3.4 (a) Reference Lena Image (I_R); and (b) Sensed Lena image (I_S) obtained by applying the RST (Rotation, Scale and Translation) transformations [6, 0.9, 4, 4] 85
3.5 The Row Variance Vectors of the Lena image (shown in Figure 3.4 (a)) at different scales factors 88
3.6 The normalised Row Variance Vectors of the Lena image (shown in Figure 3.4) at different scales 89

3.7 Row Variance Vectors of the Lena image (shown in Figure 3.4) at different rotation angles 90
3.8 Barbara reference image (top left) and sensed images obtained by applying the RST transformations (8,0.8,14,3) and with varying salt and pepper noise density (d=0:0.04:0.20) 98
3.9 Hand CT reference image (top left) and sensed images obtained by applying Rotation and Translation transformations (38,9,4) and Gaussian noise with varying variance from 0:0.01:0.1 100
3.10 Fingerprint reference image and sensed image obtained by applying Rotation and Translation transformations (17,15,-16) and Gaussian Noise with variance 0.1 101

4.1 The coloured block represents the thesis objective addressed in this chapter. 105
4.2 Lena reference and 90° rotated images 107
4.3 The Sinogram of the Lena reference and 90° rotated images (the rotated image Sinogram is the shifted version of the reference image Sinogram). 108
4.4 The block diagram of computing the rotation invariant Variance Vector 111
4.5 The process of converting an image in the pixel domain to a rotation invariant Variance Vector 111
4.6 The normalised Variance Vectors of the reference and rotated Lena images (shown in Figure 4.2) 112
4.7 Ten noise-free logo images from the UMD dataset [15] and the samples of noisy images generated by adding the salt and pepper noise with densities ranging from 0.00 to 0.20 with increments of 0.04 . . 120

4.8 Precision-Recall curves of the proposed and R-RST method at different noise densities 0.00:0.04:0.20 121
4.9 Precision-Recall of the proposed method at different noise densities 0:0.04:0.20 and with different numbers of projections 122

5.1 The coloured block represents the thesis objective addressed in this chapter. 129
5.2 Lena reference and sensed (45° rotated) images 136
5.3 (a) Square shaped Lena reference image region; (b) Square shaped rotated (45°) Lena image region; (c) Sensed image region rotated back by 45° 136
5.4 (a) Circular shaped Lena reference image region; (b) Circular shaped rotated (45°) Lena image region; (c) Sensed image region rotated back by 45° 136
5.5 The process of identifying common regions between unaligned partial images without aligning them 140
5.6 The Variance Vector of the Lena image circular region shown in Figure 5.4 (a) 141
5.7 (i) Lena reference Image (I_R); (ii) Lena sensed image (I_S) obtained by applying the rotation and translation transformations [6, 4, 4] . . 142

6.1 The coloured block represents the thesis objective addressed in this chapter. 147
6.2 Different scenarios that result in partial fingerprints (a) miniaturization of sensors; (b) a distorted fingerprint; (c) a low quality fingerprint; (d) fingerprints found at crime scenes [16]. 149
6.3 The common regions between the Lena original and rotated images and their rotation invariant Variance Vectors 160
6.4 Specifications of the FVC2002 and FVC2006 datasets 165

6.5 The effect of the number of regions on the accuracy achieved by the proposed approach on the FVC2002DB1 dataset. The accuracy is achieved by considering the first impression of each person as a sensed image. 167

6.6 The effect of multiple regions on the accuracy achieved by the proposed approach on the FVC2002DB2 dataset. The accuracy is achieved by considering the sixth impression of each person as a sensed image. 167

6.7 The effect of the size of regions on the accuracy of the proposed approach. The accuracy is computed for three different diameters (71, 101 and 141 pixels) of circular regions by considering the first impression of each person as a sensed image for the FVC2002DB1 dataset. 168

6.8 The average accuracy with respect to the number of regions on the FVC2002DB1 dataset 171

6.9 Region from a noiseless fingerprint (top left) and its transformed versions obtained by adding salt and pepper noise of varying densities d=0.04,0.08,0.012,0.16,0.20 175

Chapter 1

Introduction

1.1 Preamble

Image alignment plays a crucial role in several image processing and computer vision applications [17]. It is a process of bringing two or more images of same scene taken from different view points and at different time instances into geometrical agreement. As a result of this, the pixels in two images match the same physical region depicted in images [18]. It is the stepping stone for many applications such as image registration, object recognition, biometric recognition, image retrieval, satellite image analysis and character recognition to name a few. Practically, all systems which are used to evaluate the images require alignment of images or a closely related operation as an intermediate step [19].

In image registration based applications where post processing is done manually, alignment helps in bringing the similar points of images into correspondence and makes the evaluating process easier, whereas in recognition based applications in which alignment of images acts as an intermediate step, it helps in matching two images [14]. Medical image alignment and non-destructive testing image alignment are two examples of image registration applications, whereas object recognition, biometric recognition and character recognition are the examples of image match-

ing applications. Moreover, alignment is needed for the integration of images in different applications such as space image integration [20]. The amalgamation of images can result in complementary information and therefore improve the performance of decision making process.

In the last few decades, image capturing devices have gone through a rapid development resulting in a diverse range of obtained images. The availability of a variety of images has provoked research in image alignment. Due to the ease of capturing images, advances in data processing, networking and storage technology, the number of image matching applications has increased. These applications demand an accurate and efficient image alignment method.

Over the years a large number of methods have been developed for different types of data and problems in computer vision [21]. These methods have been separately studied for different applications, resulting in a large body of research. Predominantly, the methods can be divided into two broad categories, namely, feature based and pixel based (with or without using transforms such as Radon, Fourier or Wavelet) [9] [22]. The features of images represent high level information, hence it can be used as a measure for aligning images. However, extraction of features faces a number of difficulties such as requiring extensive image processing and extracting false features. There is an especially high probability of extracting false features from low quality images and partial images do not contain enough features. Because of these issues, the feature based techniques cannot be applied for matching images captured under different scenarios.

Pixel based techniques are suitable for different types of images irrespective of pixel to pixel correspondence between them and do not require extensive image processing like feature based techniques [3]. However, these approaches are computationally expensive [23] [10] [14]. To improve the computation time of pixel based approaches, different transforms are used which make it easier to compute the misalignment between images. Recently, a number of transform based meth-

ods have been proposed which are highly accurate [14]. However, the methods either depend on pixel-to-pixel correspondence between images or use exhaustive approaches to align images. Therefore, an accurate and efficient image alignment technique which is independent of feature and pixel-to-pixel correspondence between images is important for different applications of image processing and computer vision.

In image matching applications such as biometric or object recognition, alignment is used as an intermediate step. Several image matching methods based on alignment of images have been developed by researchers [10] [14]. However, using alignment in a matching process has many disadvantages. The main ones are: (i) the time required for aligning the images affects the efficiency of the entire matching process; (ii) most of the alignment methods rely on features of images which in turn depend on pixel-to-pixel correspondence between images; and (iii) misalignment of images can result in false matching of images. Therefore, for matching images efficiently and irrespective of the pixel-to-pixel correspondence, it is important to match images based on their pixel information and without aligning them.

To recapitulate, accurate and efficient alignment and matching of images without extracting features regardless of pixel correspondence is a challenging task. The next section introduces the issues and challenges faced in accurately and efficiently aligning and matching images.

1.2 Issues and Challenges

In computer vision and image processing, alignment of images plays a vital role [24] [22]. Alignment of images is a process of transforming one image to the coordinate system of the other image. The alignment of images has a different use in different applications such as in object recognition it helps in matching two images whereas

in image registration it brings the similar points of the images into correspondence [14]. The alignment of images seems a clear problem that should have a universal solution, however, this is still far from the state of the art. Due to an increase in the number of image matching based applications, image alignment has become a complex problem that acknowledges a variety of solutions. Over the past few decades, research in image alignment has substantially increased due to a growing availability of digital images in different computer vision applications.

The earliest method of aligning images was manual alignment in which the distinctive points in images are selected manually prior to these points being used to compute the geometrical transformations present between images. Due to its ease of implementation, this method is widely used in the remote sensing community. However, this method is laborious and prone to error. That makes it difficult to process large amounts of data [9] [22]. Therefore, researchers have developed automatic image registration methods. As discussed earlier, image alignment can be classified into different categories: feature based and pixel based. Feature based approaches identify the correspondence among features of images. Whereas, pixel based approaches use intensity patterns of pixels by means of correlation and by computing the transformation parameters between sensed and reference images [25].

Numerous researchers have worked on feature based methods in different fields such as pattern recognition, computational geometry and image processing [26]. These methods mainly consist of three steps: (i) extraction of reliable features in images; (ii) matching of features; and (iii) alignment of the images based on matched features [25]. One of the preliminary steps is the detection and extraction of features in images. This means the images are represented by a set of feature points. Generally, feature points represents the high-content information that can be represented in two directions simultaneously such as line intersection or corners. Many operators have been developed to detect features. Commonly used detectors

are Harris and Stephens [27] and Forstner and Gulch [28]. Schmid et al. [29] have evaluated a number of feature point operators for their repeatability and found that the Harris operator is the most repeatable. In 2004, Mikolajczyk and Schmid [30] have extended the Harris operator for the affine transformations. The feature points are generally application specific. In object recognition, feature points are commonly represented by corners, edges or T-points. The most common methods of detecting these features are discussed in [2] [31] [27] [32]. The Scale Invariant Feature Transform (SIFT) descriptor has been widely used for many image registration applications. In 2006, Yang et al. [33] used SIFT for the extension of the dual bootstrap Iterative Closest Point (ICP) originally developed for image registration. However, in fingerprint recognition, singularities or minutiae points are used as feature points [3]. Matched feature points are used as reference points in both images for aligning them. However, there are issues related to these approaches such as (i) accurate detection of reliable features is very difficult; (ii) detection of features is time consuming as it requires extensive image processing; (iii) images which do not have exact pixel-to-pixel correspondence such as partial images may not consist of enough features required for alignment; and (iv) application dependence. These issues complicate the process of aligning images and affect the accuracy and efficiency of the image matching process. Therefore, aligning images using pixel information of images is a better option.

Another major method of registering images is pixel based alignment that is based on the correlation of images [17]. These approaches use the smaller regions of images for matching purposes and therefore, it is applicable to all types of images regardless of images having pixel-to-pixel correspondence or not. Correlations compute the similarity between images by comparing the pixel intensities [34]. In the absence of reliable features, correlation is a trustworthy method of aligning images. The correlation between two images is greater if and only if the amount of pixel correspondence is large. For example, when correlation is used for aligning

fingerprints [10], in order to find the best pixel-to-pixel correspondence between sensed and reference images, the sensed image needs to be rotated by each possible angle and compared with the reference image. This is an exhaustive approach that makes it practically impossible for real time applications. Lucas et al. [35] have developed another pixel based alignment algorithm namely optical flow that has been used for aligning images in many applications. However, solving optical equations is a time consuming process. Although, the correlation based approach is widely used approach for image alignment, it has limitations that need to be addressed. The brute-force searching for optimised alignment is computationally expensive. The number of operations required for computing the correlations depends on the size of regions of image under consideration and number of search windows. It is acceptable for smaller sized regions. However, developing a faster pixel based alignment method is still a challenging problem. Moreover, the basic correlation measures are affected by noise and therefore, more statistically complex similarity measures are required.

Transform-based registration methods solve some of the issues of correlation based approaches. These methods align the images by computing the underlying transformation parameters. These methods are based on the premise that information in the transformed image makes the geometric transformation easier to compute. The commonly used transforms are Fourier, Wavelet and Radon [36]. These approaches compute the underlying transformation parameters using the properties of the transforms. Many approaches based on transforms have been proposed in the literature that do not require the extraction of features of images. Recently, Radon and Fourier transforms have been used by researchers for computing the transformation parameters. In [37], Fourier phase matching is used to register the computed tomography and radiography projections which has the ability to allow defect detection. However, this method is only able to find the translation between two images. Even in translation it is difficult to compute the

large displacements. Chen et al. [19] have used the Radon transform with the Fourier transform to register the images. The results of this approach are fairly less accurate on some images [14]. Mooser et al. [38] have developed a method that can deal with all degrees of freedom of transformation. The parameters are estimated as an optimization problem which uses a region based optimizer. However, the error in estimation of both parallel and fan beam geometries is very large.

Estimation of the transformation parameters relating two images using Radon projection has gained attention recently. Actually, working directly with Radon projections eliminates the need for reconstruction of images which is prone to computationally exhaustive reconstruction artefacts. A method of computing the parameters for a binary image subjected to translation, rotation, reflection and scale by using only the Radon projections is proposed in [39]. However, the objective function used to compute the rotation angle cannot be generalized for angles in the range $[0, 2\pi]$ [14]. In [14], authors have proposed a method that can find a rotation angle lying between $[0, 2\pi]$ without additional computation unlike those used in [39]. The method is highly accurate and robust to noise, but the process of estimating the parameters is computationally expensive which slows down the entire alignment process. Not only this, most of the above methods are applicable to only those images which have pixel-to-pixel correspondence. Applying these methods to the images without pixel-to-pixel correspondence such as partial images will result in inaccurate parameter estimation.

As discussed above, the existing methods either are not accurate or are computationally expensive. Therefore, developing a pixel based alignment method that computes the parameters in less computation time without affecting the accuracy and is also applicable to images which do not have pixel-to-pixel correspondence is still a challenging problem.

The alignment of images is used as an intermediate step in matching the images. Not only does the computation time needed for aligning the images affect the

overall efficiency of the matching process, but also misalignment can result in false matching of images. Moreover, the parameter based alignment methods are not applicable to partial images whereas feature based methods find it difficult to have enough features in such images. Therefore, matching all types of images efficiently and accurately without extracting features is still a challenging problem. For faster matching of images, it is better to match the images without aligning them and irrespective of the pixel-to-pixel correspondence between them. This will not only result in matching the images accurately and in less computation time, but will also match the images whether the images have pixel-to-pixel correspondence or not.

1.3 Motivation

Due to the rapid increase in the number of images, a large number of image matching applications have been developed. One of the recent applications is unlocking iphones using fingerprints or face images that needs accurate and effiicent matching of images. The motivation of the thesis is derived from this increasing demand for accurate and efficient image alignment and matching in various applications. The false matching or misalignment of images can result in severe consequences. Some of the best examples of image registration where high accuracy is demanded are alignment of space images in astrophotography, brain image alignment in neurosurgery where an error of even 1^o can result in wrong star localisation or wrong tumor spatial limits, respectively [14].

For personal identity authentication these days, not only electronic devices but many large organisations have started using biometrics. The false matching of images can cause loss of private information or breaches in security. On the other hand, wrong matching of images in applications such as criminal identification can destroy someone's life. In 2004, a lawyer named Brandon Mayfield was falsely

convicted for Madrid train bombings. At the crime scene, Spanish authorities found some fingerprints on a plastic bag and these fingerprints were matched to the FBI database. To one of those fingerprints 20 possible matches were found in the database and the top match was that of Brandon Mayfield. The FBI declared the match as a 100% match to his fingerprints, which were taken as a standard procedure when he joined Army, and arrested him. This happened due to mismatching of images because the fingerprints found at the crime scene were unclear and partial. In many other applications, images are captured in uncontrolled environment such as images captured by smart phones. The images captured by smart phones can be unaligned. Dirt on the camera can distort the images. These issues of images increase the complexity of the matching process and result in mismatching of images.

Generally, the mismatching of images results due to noise present in images and the absence of required pixel correspondence between images. Accurate and efficient matching of such images by manual inspection and existing feature based automatic image matching methods is quite difficult. This is because the accurate manual inspection of a large number of images is tiresome and time consuming. Also, using the automatic existing image matching techniques is not accurate as they rely on extracting features. Different features are extracted for different types of applications. For example, object recognition techniques extract the interesting feature points such as corners, edges, t-points or crosses in the images, whereas fingerprint recognition techniques use minutiae, singularities or pores present in fingerprints. The extraction of these feature points requires extensive image processing which is time consuming and affects the efficiency of matching process. Moreover, due to the absence of required pixel-to-pixel correspondence in images most of the existing techniques do not work. In addition, when images do not have high pixel-to-pixel correspondence then relying only on a few features of images is not enough. These issues motivate us to use pixel based information of images

which does not require extensive image processing and is independent of applications. Hence, a pixel based matching method is required which can accurately and efficiently match the images without extracting features and without bothering about pixel correspondence and noise present in images.

1.4 Research Objectives

The current techniques face challenges in matching partial images which do not have exact pixel-to-pixel correspondence. Either the methods do not perform well in terms of accuracy or are computationally expensive. Therefore, as discussed earlier, to deal these issues a pixel based technique is a better approach. This thesis overcomes the above limitations by proposing a pixel based method that can efficiently align and match images without extracting the features with a performance better than the existing techniques. It is applicable to images irrespective of the pixel-to-pixel correspondence present between them. In order to achieve the aim of this research, two objectives have been formulated as follows:

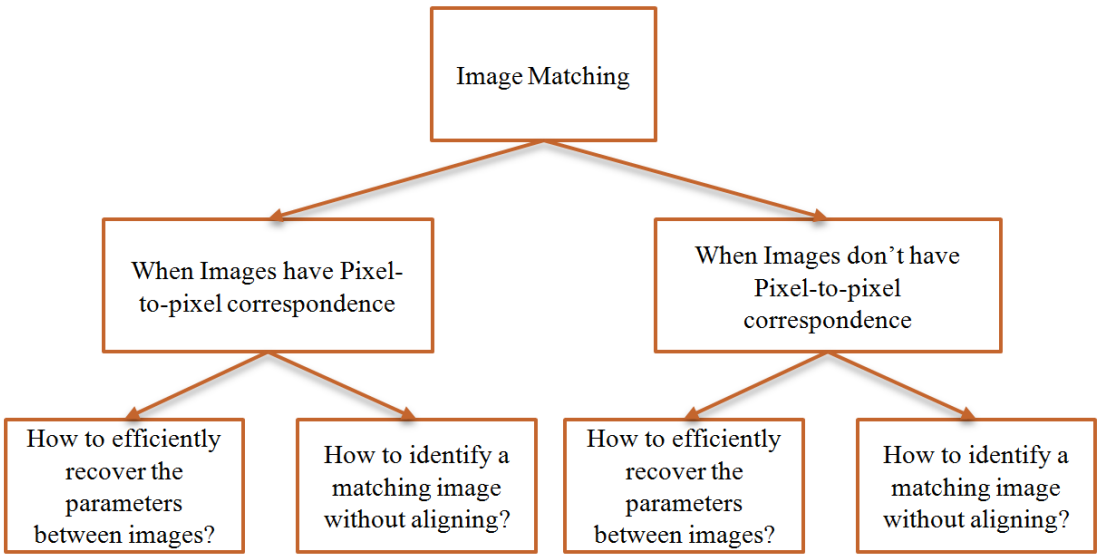


Figure 1.1: The block diagram of research objectives

i) Compute the transformation parameters between images to accurately and efficiently align them and match the aligned images.

ii) Develop a image matching method that can accurately and efficiently match images without aligning them.

The block diagram of research objectives is shown in Figure 1.1. For matching images without needing to align them, the existing similarity metrics cannot be used as they need aligned images, therefore, a new similarity metric is required that can compute the similarity between unaligned images. On the other side, in the case of partial image alignment and matching, a number of other issues are faced as they have only a part in common between them. The issues are:

(a) How to efficiently extract common regions between images without aligning them?

(b) What should be an appropriate shape of regions to have pixel-to-pixel correspondence?

(c) How to compute the similarity between unaligned partial images by using most of the given information?

1.5 Research Contributions

The contributions of the thesis can be summarized as follows:

i) The method of computing the transformation parameters is divided into two categories based on the pixel-to-pixel correspondence present between them:

   (a) In order to accurately and efficiently align the images without extracting features, the thesis has developed a transform based method to compute the transformation parameters between images. Subsequently, to deal

with noise present in images, the robustness of the proposed method to noise is also verified. Moreover, to substantiate the application independence of the proposed approach, the method is evaluated on images of different applications. However, this approach is applicable only when images have pixel-to-pixel correspondence.

(b) In order to align the images which do not have pixel-to-pixel correspondence, a region based approach is developed. For that, a method to extract the common region between unaligned images and an appropriate shape of the region for extracting a common region is proposed. For efficient detection of a common region, an optimization technique is also proposed. The extracted common regions can be used for aligning the images which do not have pixel-to-pixel correspondence where most of the existing methods fail to perform.

ii) The method of alignment-free matching of images is classified into two categories based on pixel-to-pixel correspondence present between them:

(a) In order to match the images accurately and efficiently, an alignment-free transform based method is proposed for images having pixel-to-pixel correspondence that can match the images without computing all the parameters and without extracting features. As images are matched in the transformed domain and without using features, a new similarity metric is proposed to compute the similarity between images. In addition, a method to further improve the efficiency of the system is presented. The effectiveness of the proposed method is verified by testing on logo datasets.

(b) In order to match the images that do not have pixel-to-pixel correspondence, a region based alignment-free method is proposed. For matching

images, a method of computing global similarity based on local similarities is proposed. Moreover, the effect of multiple regions on accuracy is studied. Results show that overlapping regions can improve the accuracy. The effectiveness of the proposed method is studied by applying to standard fingerprint datasets. Its robustness to noise is also demonstrated.

1.6 Thesis Organization

The rest of the thesis is organised as follows:

Chapter 2 introduces the fundamentals of image matching from the perspective of the thesis, and provides a review of relevant literature in the area of image registration and matching. Different approaches to improve the image registration and matching in different applications are discussed in this chapter. The chapter demonstrates the challenges in aligning and matching images regardless of pixel-to-pixel correspondence.

Chapter 3 discusses the theoretical foundation of computing the parameters used for aligning images having pixel-to-pixel correspondence. The effectiveness of the method is tested on images of varying classes. Results are compared with the state of the art methods. Also, the robustness of the proposed method is tested with respect to different types of noises.

Chapter 4 provides the theory of alignment-free transform based image matching methods for matching images having pixel-to-pixel correspondence. A similarity metric is developed that can compute similarity between images in the transformed domain. The chapter includes discussions and evaluation of the results of the proposed technique by running experiments on logo datasets. Results are compared with the existing techniques. In addition, robustness of the proposed

method to noise is verified. A method to further improve the efficiency is also proposed.

Chapter 5 describes a method to align the images which do not have pixel-to-pixel correspondence. The theory of extracting the common information (common region) between images and the appropriate shape of the region is discussed. Performance enhancements are proposed to optimize the process of extracting the common region.

Chapter 6 discusses a region based alignment-free matching method for matching images which do not have pixel-to-pixel correspondence. The methods of using all the given image information is proposed. The method is tested on standard fingerprint datasets. The chapter also includes strategies of optimizing the image matching method.

Finally Chapter 7, summarizes the contributions made by the thesis and discusses possible future research directions.

Chapter 2

Background and Related Work

2.1 Preamble

Chapter 1 argued that accurate and efficient alignment and matching of images is essential for different applications. The aim of this chapter is to review existing techniques and discuss the main issues and concepts relating to the improvement of the alignment and matching of images. Since the aim of this thesis is to improve the alignment and matching accuracy and efficiency, evaluating existing techniques is very important.

Depending on the features used for aligning and matching images, we have divided the existing techniques into two main categories namely, feature based and pixel based techniques. We argue that current techniques fail to accurately and efficiently align and match images. The full or partial natures of the images considered signify that the images can or cannot have pixel-to-pixel correspondence respectively.

Let us suppose two images I_S (sensed[1]) and I_R (registered) are provided for aligning or matching. For aligning images, either the features of images are used or similarity transformation parameters are computed. However, both of these

[1]sensed or transformed are used interchangeably

methods either are computationally exhaustive or not applicable to partial images. On the other hand, in an image matching process a similarity score is computed. The score determines whether the images are matching or not. However, before computing the similarity score, the images need to be aligned. The alignment process may affect the accuracy and efficiency of the matching process. In addition, due to lack of information present in partial images, the accuracy of matching can also be affected which is essential in recognition processes.

The chapter has been structured as follows. In Section 2.2, the general image matching process and in Section 2.3 categories of matching processes are discussed to help understand the following discussions. In Section 2.4, the current feature based techniques are discussed. Furthermore, the pixel based techniques are discussed in Section 2.5. There are many existing techniques which use both types of features. The most relevant ones are discussed in Section 2.6. Learning based techniques are discussed in Section 2.7. Following the hybrid techniques, a comparison of existing techniques is discussed in Section 2.8. Finally, in Section 2.9, the conclusion of this chapter is presented.

2.2 Image Matching Systems

In this section, the general image matching system is discussed. Image matching is the research foundation of different computer vision applications such as object recognition, biometric recognition, image registration, 3D reconstructions, etc [40]. It has wide spread use in medical imaging, remote sensing, artificial intelligence, etc. In an image matching process, the alignment of images is a determining step (as shown in Figure 2.1), as alignment simplifies the matching process. The motive of image alignment is to find the best possible transformation between images that best aligns the structures in input images. One of the causes of misalignment in images is comprised of the rotation and translation differences between them. The

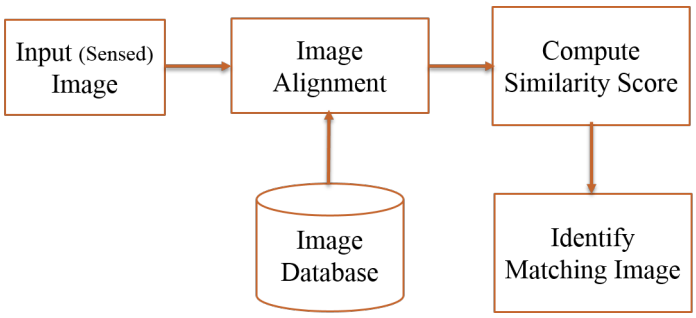


Figure 2.1: The block diagram of an image matching process

perfect alignment of images can eliminate the variations present in images (such as rotational and translational difference) and greatly improve the efficiency and accuracy of an image matching system [3]. Alignment of images is required in different applications which can be broadly divided into two types: (i) registration applications in which images need to be aligned, where a final decision is made manually on aligned images; and (ii) recognition applications where images are aligned and a similarity score computed to find the matched image.

The patch based translational alignment method developed by Lucas and Kanade in [35] is an early example of a widely used image alignment method. Variants of this method have been used in different motion-compensated video compression schemes [41]. Applications such as video summarization, video compression and video stabalization are based on similar parametric motion estimation techniques [42][43][44]. In the mid 1990s, alignment methods started being used for the construction of wide-angle seamless panoramas from regular hand-held cameras [45][46]. The recent work in this area has directed the computation of similarity transformation parameters between the images for computing the parameters.

The methods of computing parameters between images can be classified into two categories: feature based matching and pixel based matching. The feature based matching techniques use the extracted features of images to align them, whereas, pixel based matching techniques use the gray scale information of images [47][48]. The existing feature based techniques use the extracted features such as

corners, edges, minutiae etc. for alignment of images [49]. However, detection of these features accurately requires high quality images that are not common in real life. In addition, the acquired images may be partial. The fragments might not have enough required features for alignment. Hence, these techniques will either not be applicable or result in incorrect alignment. On the other hand, pixel based approaches can align the images accurately without using any specific feature [10]. However, these approaches use brute-force searching for aligning images which is very time consuming. These techniques try to minimize pixel-to-pixel dissimilarities between images, whereas feature based techniques extract the image specific features and then match the images [50][51][52]. The biggest benefit of using a feature based approach is that these techniques have the ability to automatically find the overlap existing between images.

As shown in Figure 2.1, image matching is a process of comparing two images and computing a similarity score between them. Once the images are aligned, the similarity score is computed between images and based on the similarity score the matching image is identified. Many image matching techniques have been proposed in literature (discussed in Section 2.3). Due to noise and other variations in the images, it is not possible that two images match exactly. Therefore, similarity between two images is measured in terms of a matching score where a higher matching score represents greater certainty that two images are similar.

2.3 Categorization of Image Matching Techniques

The previous section discussed that alignment of images is a basic requirement for matching images. The main idea of the alignment process is to search iteratively for the geometrical transformation that when applied to the transformed image optimize (minimize or maximize) the similarity metric. One of the oldest and most widely used application of alignment in computer vision is the seamless photo-

mosaicing to make beautiful ultra wide-angle panoramas [9]. These days, it is used in other computer vision applications such as texture analysis, medical image diagnosis, object recognition, image retrieval and so on. In the past decades, the research into image registration has had enormous growth.

Manual alignment of the images is the earliest technique of aligning images. For that, distinctive points are selected manually and these points are used for finding the correspondence between images. For example, in manual fingerprint identification, human experts examine the fingerprints starting from level 1 features (such as minutiae) to level 3 features (known as lower level features) [3].

Subsequently, automatic image matching methods were developed. These were inspired by manual procedures of image matching. However, there are many other techniques which are specially developed for automatic matching of images. Most of the existing techniques do not face any difficulty in matching good quality images. However, the matching of partial and poor quality images reliably is still an unsolved problem. For instance, the coarse evaluation of the existing fingerprint matching techniques on a $FVC2000$ database shows that 80% of the false rejection errors are made by 20% of the low quality fingerprints present in the database [3].

Matching methodologies can be classified using feature space image information. This information can be intensity of the pixels, intensity gradients or the structures extracted from the images such as edges, corners etc. Methodologies based on the pixel intensity are called intensity (pixel) based and those based on geometrical structures are termed as feature based [8]. The advantages and limitations of each approach are discussed below.

2.4 Feature based Techniques

Feature based matching techniques are by far the most commonly used and researched techniques in image matching. These approaches have been used since

the early days of stereo matching [53] [54]. Nowadays it is used for different applications such as image stitching, image registration, biometric recogntion amongest other [52] [51]. It is a mature approach that has been used for years and is based on extraction of the salient features of images.

Significant lines (roads, rivers, region boundaries), point (region corners, points on high curvature, line intersections) or significant regions (lakes, fields, buildings or shadows) are generally considered as features in images. In most cases, the main algorithms of feature detection take the line intersections or centroids of regions segmented from an image or the local maxima of Wavelet transforms as feature points. Some of the famous line detectors are the Canny edge detector [55] and the Laplacian of a Gaussian based detector [56]. The regions are often represented by the weighted averages of pixel values which are invariant to the geometrical transformations [57]. Moreover, features need to be distinctive, easily extractable, uniformly distributed in images and invariant with respect to time [2] [31].

In feature based techniques, first the features of images are extracted (discussed in Section 2.4.1), second these features are used to find geometrical transformation parameters between images to align them (discussed in Section 2.4.2) and finally matching is performed by computing similarity scores (discussed in Section 2.4.3).

2.4.1 Feature Detection

For matching images, different types of features can be extracted from images depending on the applications. For example, in object recognition, generally corners or edges are extracted from images [9], whereas in fingerprint recognition minutiae and singularity points are extracted [58]. Many feature detectors have been developed in the literature and some of them are discussed below.

Feature points are the well defined positions in images that can be detected robustly. Repeatability is the most desirable property of detectors so that detectors

can detect the same features in transformed (sensed) images as well. Feature detection is actually used as an initial step in many computer vision applications. One of the earliest detectors, developed in 1983, extracts corner like features in images [59]. This corner detector was modified in [27]. Since then the Harris detector is widely used for image matching tasks. The detector is based on the eigenvalues of the Harris matrix which is given by:

$$A = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix} \tag{2.1}$$

where I_x and I_y are the partial derivatives of image I in x and y direction respectively. The angle bracket notation takes the average over a patch around the current pixel into consideration. For a point to be an interesting point, both eigenvalues of the Harris matrix at that point should be large [27]. Later, a technique was proposed that used a combination of translational and affine-based patch alignment for tracking the feature points through an image sequence [60]. The feature points invariant to scale were proposed in [2] and [30] which are very helpful for matching images captured at different scales. The easy way to get scale invariance is to find the scale-space maxima of the Difference of Gaussian (DoG) computed on a sub-octave pyramid [61].

FAST (Features from Accelerated Segment Test) is a robust corner detector which performs faster than most existing detectors for real time applications, but it is not scale invariant [1]. In this detector, for a point p to be a feature point, consider a circle of 16 pixels (represented by x) around this point and select a number n (originally 12) and threshold t (shown in Figure 2.2). If n contiguous pixels of a circle are brighter than $I_p + t$ or darker than $I_p - t$ where I_p is the intensity of point p , then p is a corner. This method does not reject a sufficient number of candidates for $n < 12$. Multiple features are detected near to each other.

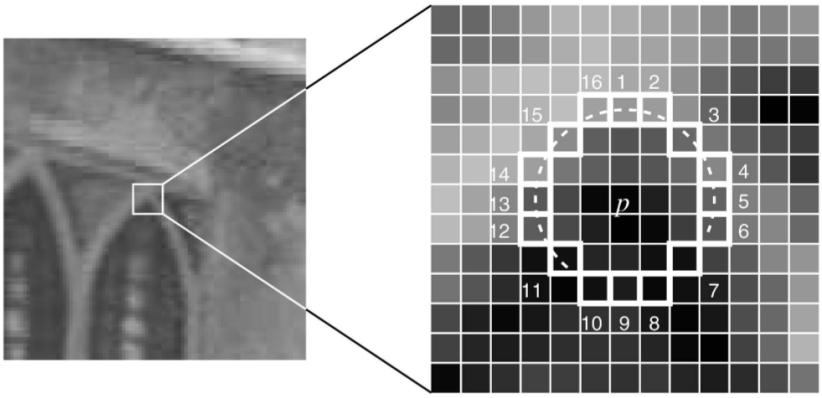


Figure 2.2: Corner detection in an image patch by using a twelve-point segment test [1]

This problem is solved by learning a corner detector in which the ID3 algorithm [62] is used to find a decision tree. The tree is used for fast detection of features in test images. To remove the problem of detecting feature points nearer to each other, a score function V is calculated as given in Equation 2.2.

$$V = max\left(\sum_{x \in S_{bright}} |I_{p \to x} - I_p| - t, \sum_{x \in S_{dark}} |I_p - I_{p \to x}| - t\right) \tag{2.2}$$

where, S_{bright} and S_{dark} are the sets of brighter and darker pixels, $p \to x$ represents the pixel at position x relative to p . The feature points with high values of V are kept and the others are ignored. Corners are not the only type of features that can be used for aligning the images. To estimate the similarities between the images, line segments can also be used [63]. These detectors consider the points at edges as feature points, because at edges the brightness of an image changes abruptly. Edge detection can be of two types, search based and zero crossing based [64]. A search based method computes the edge strength that is generally the magnitude of the first order derivative and after which local directional maxima are found. In a zero crossing based approach a point is detected as interesting point if the second order derivative at that point is zero. Sobel [65], Marr Hildreth [56] and Canny [55] detectors are examples of widely used edge detectors. The Canny edge

detector [55] is one of the famous search based detectors in which non-maximal suppression is used for finding the local maxima. Non-maximal suppression is a method of eliminating the points that are not lying on edges. Finally, hysteresis thresholding was used for finding the real edges. Hysteresis thresholding uses two thresholds, maximum threshold and minimum threshold. Edges having a gradient magnitude more than the maximum threshold are considered as real edges; edges having a magnitude less than the minimum threshold are discarded. But if the magnitude lies between both thresholds then an edge is selected or discarded on the basis of connectivity, which means if the edge is connected to the real edge then it is considered as an edge otherwise it is discarded.

Affine invariant regions or salient regions are also used to find the correspondence between images in [66] [67]. Blobs are the regions in an image which are different from other parts of the image either in brightness or colour. Laplacian of Gaussian (LoG), Difference of Gaussian (DoG) and the determinant of the Hessian are the famous blob detectors. Parker and Crowley [68] used scale space and detected peaks and ridges by changing the scales. Later, Lindeberg [69] studied the need of proper scale. Lowe [2] invented the Scale Invariant Feature Transform (SIFT) descriptors which detect the scale and rotation invariant features of an image. SIFT uses DOG which is calculated by simply subtracting the Gaussian smoothed images at different scales as shown in Figure 2.3. For finding the local maxima or minima, non maximal suppression is used. Non maximal suppression compares a pixel with 26 other pixels (8 from DoG of the same scale, 9 from DoG of a larger scale and 9 from DoG of a smaller scale). If the pixel is maximum or minimum among all the pixels then it is considered as a local maximum or local minimum.

Later, Bay et al. [31] have developed the Speeded up Robust Feature (SURF) descriptor which uses a Fast Hessian detector on an integral image. The integral image and the Hessian matrix enhance the accuracy and decrease the computation

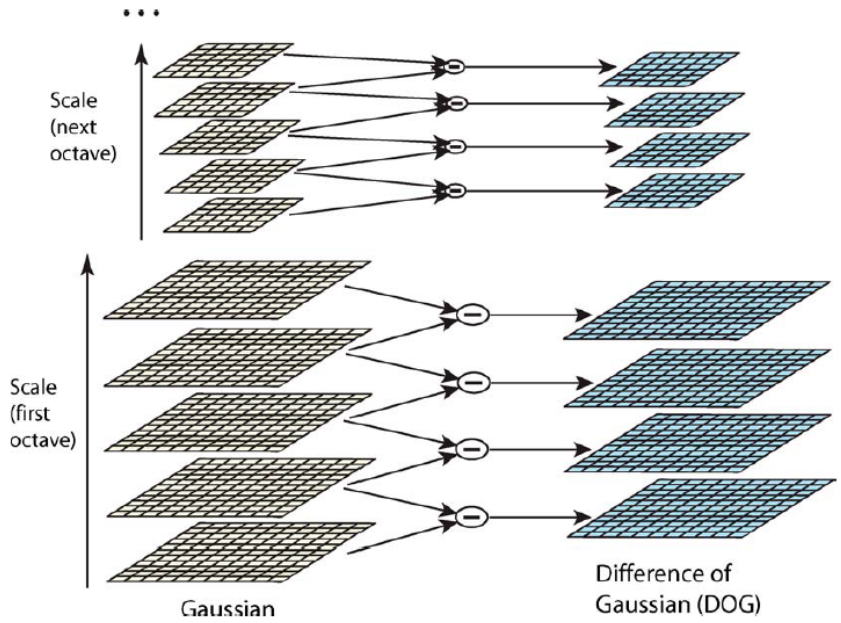


Figure 2.3: In each octave the same image is convolved with a Gaussian of increasing scale, and finally these smoothed images are subtracted for producing a Difference-of-Gaussian [2].

time. The integral image $I_{\sum}(X)$ at a point X=(x,y) is

$$I_{\sum}(X) = \sum_{i=1}^{i \leq x} \sum_{j=1}^{j \leq y} I(i,j) \tag{2.3}$$

where I(X)is the input image. The Hessian matrix at a given point $X = (x, y)$ is

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \tag{2.4}$$

where $L_{xx}(X,\sigma)$ is the convolution of the integral image with the second order derivative of a Gaussian $\frac{\partial^2}{\partial x^2} g(\sigma)$ at point X and scale σ. $L_{xy}(X,\sigma)$ and $L_{yy}(X,\sigma)$ are similarly defined. For finding accurate feature points, all pixels where the determinant of the Hessian matrix is less than the threshold value are rejected [31]. After thresholding, local minima or maxima are found by non-maximal suppression. In SIFT [2] and SURF [31] responses for all pixels at all scales are not calculated, but this is not the case with CenSurE (Center Surrounded Extremas)

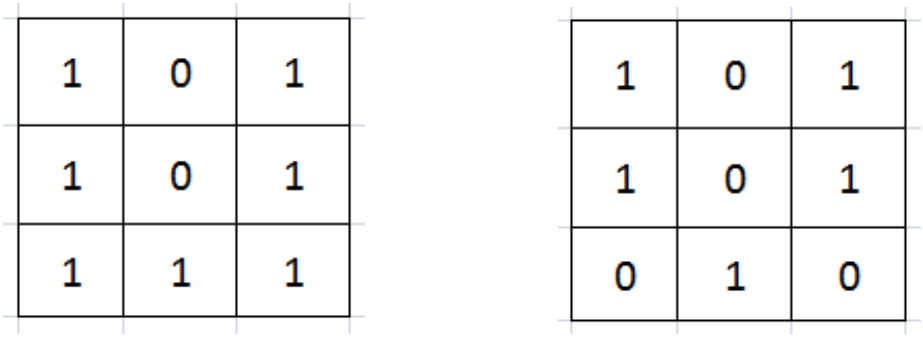


Figure 2.4: Ridge ending and bifurcation demonstrated in terms of binary pixels

[70]. CenSurE calculates responses at all scales and all locations, while maintaining the accuracy.

Another important application that matches images by extracting features is fingerprint identification. There are numerous types of features in fingerprints, but all are not commonly used because of their uneven distribution in the fingerprint. The most commonly used feature is a minutia. The common minutiae are ridge endings or ridge bifurcations [3] as shown in the Figure 2.4. Minutiae can be extracted from a binarized or gray scale image [3]. There have been many minutiae extraction techniques such as chaincode [71], ridge flow and local pixel analysis [72], crossing number [73] to name a few. However, the simplest method of checking a pixel to be a minutia is using the 3×3 window around that pixel (shown in Figure 2.4). If the central pixel is 0 then check the other 8 pixel values in the window. If the central 0 is accompanied by only one more 0 in the window then this pixel is a ridge ending whereas if the central 0 is accompanied by 3 more zeros then this pixel is ridge bifurcation [3].

A minutia is considered a triplet $m = (x, y, \theta)$ where (x, y) represents the location of a minutia and θ represents the angle of the minutia. Let FV_R and FV_S be the minutiae feature vectors for the reference and sensed images, respectively. Each component of the feature vector represents a minutia, however, minutiae can be represented by different attributes such as location, angle or type of minutiae. The most commonly used attributes are location and angle (as shown in Figure 2.5).

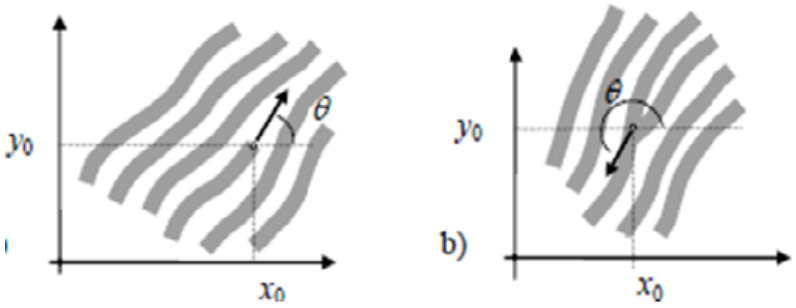


Figure 2.5: The representation of a minutia (ridge ending and ridge bifurcation) [3]

Let r and s represent the number of minutiae in FV_R and FV_S respectively.

$$FV_R = m_1, m_2, ..., m_i, ..., m_r, where, m_i = x_i, y_i, \theta_i, for \quad i = 1, 2, ...r \tag{2.5}$$

$$FV_S = m'_1, m'_2, ..., m_j, ..., m'_s, where, m'_j = x'_j, y'_j, \theta'_j, for \quad j = 1, 2, ...s \tag{2.6}$$

A minutia m'_j in the sensed image is supposed to match with minutia m_i in the reference image if and only if:

$$dist(m'_j, m_i) = sqrt((x'_j - x_i)^2 + (y'_j - y_i)^2) \leq r_0 \tag{2.7}$$

$$angle(m'_j, m_i) = |\theta'_j - \theta_i| \leq \theta_0 \tag{2.8}$$

2.4.2 Feature Matching

Once the features are detected in images, the next step is to match the sensed image to the reference image. Many methods have been developed in the literature. The important ones are discussed here.

Initially discussed are spatial relationship based methods that use metrics such as distance between feature points for matching. These are usually used when the locations of detected features are ambiguous or if the neighbourhood of feature points is locally distorted. In [74], a graph matching algorithm is used for regis-

tering the images in which transformation parameters with the highest scores are set as a valid estimate. Another method of matching the feature points is a clustering technique which tries to match points connected by edges or line segments [75]. In both the sensed and reference images, for every pair of feature points, the transformation parameters that map points to each other are computed and represented as a point in the space of transform parameters. The transformation parameters that closely map the highest number of features tend to form a cluster, whereas mismatching fills the parameter space randomly. The centroid of the cluster is used to represent the most probable value of matching parameters. In [76], Chamfer matching is developed for each line feature detected in images which are matched by means of the minimization of generalised distance.

In [77], a relaxation approach based on feature points is proposed in which the confidence level of each corresponding feature pair is adjusted based on the consistency with other pairs. The confidence level is modified by an iterative approach until a certain criterion is met. This approach is quite accurate and robust as regards the non-linearities of the image, but the iterative nature makes this approach very slow and not suitable for most applications. Ratha et al. [78] used a quadruplet transformation space (dx, dy, θ, s) that is translated along the x and y directions, rotation (θ) and scale (s) parameter, respectively. The finite sets of all possible values for 4 components are pre-defined to reduce the computational complexity. To find the transformation parameters between two images, each possible transformation is tried on each pair of feature points and is determined by:

$$(\delta^* x, \delta^* y, \theta^*, s^*) = arg ax_{i,j,k,l} A[\delta_i x, \delta_j y, \theta_k, s_l] \tag{2.9}$$

Jiang and Yau [79] used the local structure based matching by deploying two neighbouring features around the central feature. Many parameters of two neigh-

bouring features relative to the central features are noted such as angle, distance, ridge count etc. The vector of feature point m_i that has m_j and m_k as nearest neighbours is represented by

$$V_i = [d_{ij}, d_{ik}, \theta_{ij}, \theta_{ik}, \phi_{ij}, \phi_{ik}, t_i, t_j, t_k, n_{ij}, n_{ik}] \tag{2.10}$$

where, d_{ab} represents the distance between the two feature points, ϕ_{ab} represents the angle difference between the angle θ_a of feature m_a and the direction of the edge between m_a and m_b, t_a represents the type of feature m_a and n_{ab} represents the ridge count between m_a and m_b. A weighted distance is computed between the vectors of each possible feature pair. The best matching pair is used for aligning the rest of the features. The feature vectors of the aligned pairs are compared to get a final score. This whole process is followed by a consolidation step to check if the locally matched features also match at the global level. As the local structures are not much affected by non-linear distortion, these techniques are more robust as regards non-linear distortion. However, it overlooks global spatial relationships which are highly distinctive. Therefore, the amount of discriminating information is reduced and that can cause the match of local structures of two inter-class images. To deal with this problem, an additional step of consolidation is required.

Chikkerur et al. [4] have developed a rotation and translation invariant method called K-plets. The K-plets consist of K other features along with central feature (as shown in Figure 2.6). Each feature in the neighbourhood was represented by three parameters $(\theta_{ij}, \phi_{ij}, r_{ij})$ where θ_{ij} represents the relative orientation of feature m_j with respect to feature m_i. ϕ_{ij} is the angle represented by line joining m_j to m_i with respect to the orientation of the central feature. r_{ij} represents the Euclidean distance between features m_i and m_j.

This method represents a fingerprint by a directed graph $G(V, E)$ where V is the set of the vertices representing all the features of image. E contains all the pos-

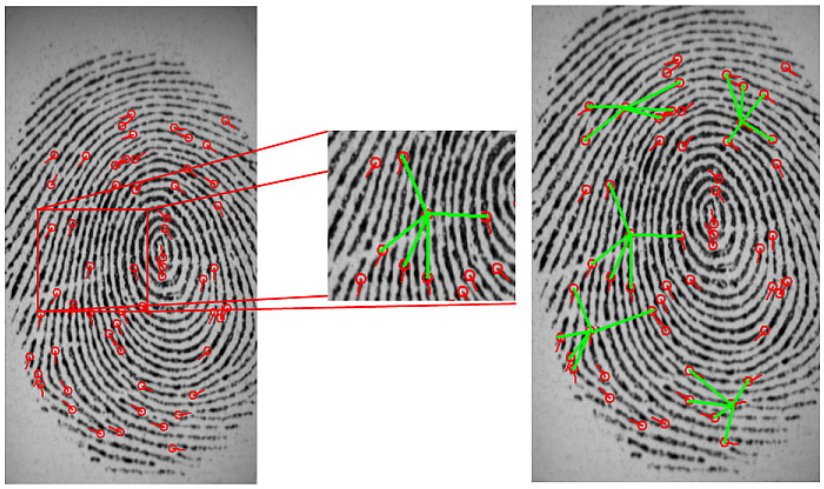


Figure 2.6: An illustration of the K-plets in a fingerprint [4].

sible edges between two features where the features are neighbours to each other. Each vertex in set V is represented by 4-parameters $(x_v, y_v, \theta_v, t_v)$ representing co-ordinates, angle and type of feature. Each edge is defined by three parameters $(r_{ij}, \phi_{ij}, \theta_{ij})$. The matching algorithm is based on local matching by matching to all the minutiae in the neighbourhood individually. The minutiae in the K-plet are organised in increasing order of radial distances. Local matching is done by dynamic programming. Coupled Breadth First Search (CBFS) is used for all possible pairs of minutiae for consolidation of local matches. The pair that results in the maximum number of matches is used for computing the final scores.

In 2009, Wang et al. [80] developed a reliable and fast fingerprint classification based on singularity points and orientation fields. The orientation field not only helps in filtering the noises in images but also represents the basic structure of image features more precisely. The method results in 96.1% of classification accuracy. The accuracy depends on the correct detection of singularity points. Moreover, the method might not be applicable to partial images as there is high probability that partial images do not have singularity points. Similarly, another directional information based classification method is proposed in [81] which results in low accuracy. Guo et al. proposed a classification method similar to the

method proposed in [80], but resulted in an accuracy of 92.7%. The method not only performs poorly in terms of accuracy but also suffers from the problem of singularity point detection.

Gao et al. [82] have developed a method in which each feature is represented by a fixed length vector. The vector of a feature of the sensed image is used to find the matching features in the reference image. This method has the lowest EER among all the feature based methods in fingerprint matching. Ratha et al. [83] used a star representation to represent feature adjacency graphs. The star associated with any feature m_j is the graph $G_j = (V_j, E_j)$ that consists of a set of vertices E representing the features that exist within a predefined threshold and the set of edges E_j between the central feature and all other features in the graph. Each edge e_{ij} is represented by 5 parameters $(m_i, m_j, d_{ij}, \phi_{ij}, rc_{ij})$ where, rc_{ij} represents the ridge count between the features m_i and m_j, ϕ_{ij} represents the angle between the edge of two features with the x-axis and d_{ij} represents the distance between two minutiae. During matching each star of the sensed image is matched with each star of the reference. The matching is done in clockwise order by traversing the graphs in increasing order of radial angle. The matching process results in a set of best matching pairs which are further checked in the consolidation step for consistency. A pair is consistent, if its spatial relation with a minimum number of other pairs in the top list is consistent.

Minutiae Cylinder Code (MCC) is a local minutiae structure based matching method which combines the benefits of nearest neighbour and fixed radius techniques and results in a fixed length code [5]. In this technique, a $3D$ structure is built from the minutiae distances and angles (as shown in Figure 2.7). It can tolerate missing or spurious features, so this approach is good for matching noisy images. The main advantage of MCC is that it has taken care of the minutiae on the borders. This is not done in other local minutiae structure based matching

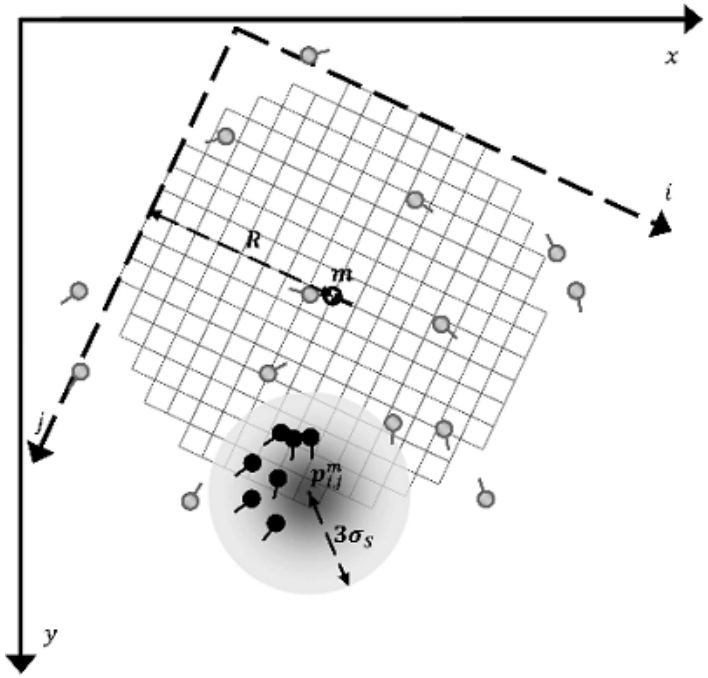


Figure 2.7: A section of the cylinder around a minutia 'm'. All the minutiae involved in the construction of cylinder are shown [5].

techniques. However, it performs poorly for the partial images that do not have enough feature points [5].

The location associated with an extracted feature point might not be the actual location of the feature. It is a better option to compare a small patch of intensities around the feature point, called a descriptor, instead of comparing only feature points [9] [84]. Using only feature points may result in more accurate scores by using incremental motion refinement. However, it is time consuming and can sometimes affect the performance [85]. Moreover, the feature descriptors are a core step of image matching and feature extraction. Over the years, many researchers in the computer application fields have worked on feature descriptors [86] [87]. The descriptors should fulfill some conditions: uniqueness (two feature points must have different descriptors), invariance (the descriptors of feature points should be invariant to corresponding transformation present between images), independence (elements of feature descriptors should be functionally independent) and stability (the feature descriptor of a deformed feature should be close to original feature

descriptor) [17]. Features resulting in the most similar invariant descriptors are considered as corresponding. The choice of type of invariant descriptors depend on feature characteristics and assumed geometric transformations. For searching the best possible matching descriptor, the minimum distance rule can be applied to the descriptions. In [60], descriptors are compared using a translational model between neighbouring frames. Then computed location estimation at this stage is used to initialize registration between patches of two images. This method is used to deal with translation differences between images. However, for handling in-plane rotation between images, a dominant orientation of each feature point is computed.

The simplest feature description is the image intensity around the feature point. For estimating the feature correspondence, the correlation between these neighbourhoods is computed [88]. Another way of computing the correspondence is by estimating the illumination direction to compensate for rotation between images and then using the coarse-to-fine correlation based registration [89].

In [6], a feature descriptor is developed based on the region around the feature (as shown in the Figure 2.8). The descriptor depends on the direction of the feature and ridge information around the feature. The descriptor is invariant to rotation and translation. Therefore, it does not require pre-alignment of images. Another advantage of this technique is that it is independent of other features as compared to the techniques discussed above. The matching score is calculated based on the similarity between the descriptors of two features.

Figure 2.9 shows a famous and widely used feature descriptor called the SIFT descriptor [2] [90]. This descriptor has mainly two issues that need to be addressed: (i) feature descriptors are based on gradient accumulations which have redundancies and are not compact; and (ii) for one local region multiple orientations are considered, which is not good for memory efficiency. Many modifications have been suggested for SIFT as the vector is quite long and it is not completely affine

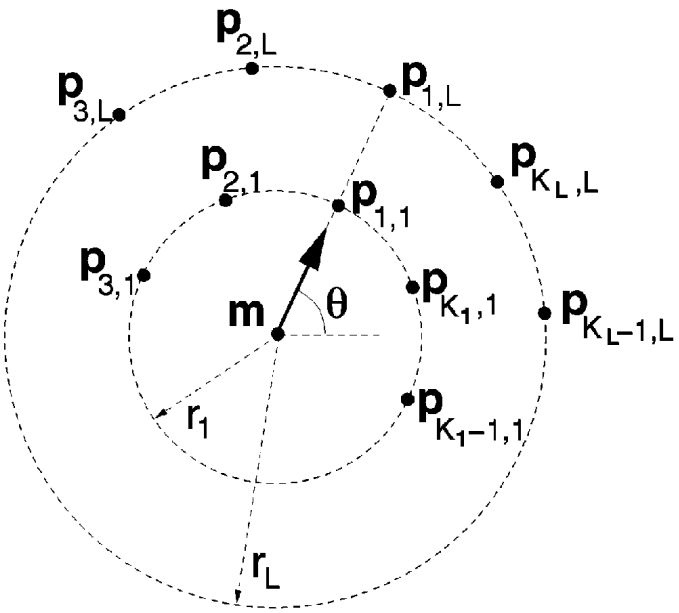


Figure 2.8: Sampling points arranged in a circular pattern around a minutia [6].

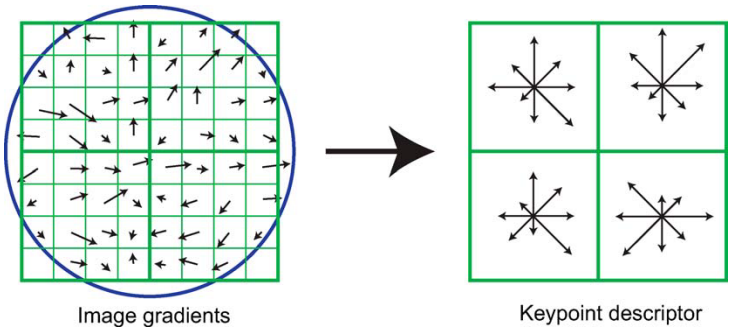


Figure 2.9: A keypoint scale invariant feature transform (SIFT) descriptor [2].

invariant. In 2004, Ke and Sukthankar developed Principal Component Analysis SIFT ($PCA - SIFT$) which results in a 36 dimensional vector [91]. Mikolajczyk et al. have proved that $PCA - SIFT$ is faster but less distinctive than SIFT [92]. Later, a descriptor called GLOH (Gradient Location and Orientation Histogram) was proposed that is more distinctive than SIFT with the same dimensions of vector but computationally more expensive [92]. Se et al. [93] used Field Programmable Gate Array (FPGA) for improving the speed of SIFT. The speed increased remarkably but the length of the vector is still a drawback of SIFT. [94] pointed out that SIFT orientations were not precise and proposed to compute the orientation at every pixel inside a local region. However, this has increased the computational complexity of the system. Texture LBP (local binary patterns) of a local region are represented by a sequence of 0s and 1s [95]. The rotation of images is identified by rotation in the binary sequences. This approach is either computationally exhaustive or replaces the gradients with other patterns to compute the orientation. To overcome the issues with SIFT, Wang et al. [96] proposed a DCT (discrete cosine transform) inspired feature transform (DIFT). DCT is used to find a dominant direction for multi-orientation in a local region. The experimental results of DIFT show that for image retrieval it achieves performance on a par as compared to SIFT, but with only 60% of the features. In addition, it reduces the dimensions from 128 to 32 and enhances precision.

In 2006, a descriptor named SURF was proposed that outperformed the existing descriptors both in accuracy as well as speed, but was not fully affine invariant or fast enough [31]. Tola et al. [97] proposed another invariant feature descriptor namely, the DAISY descriptor that is deployed for stereo matching. The performance of the DAISY descriptor is relatively good when compared to other descriptors. The classic SURF algorithm shows unstable performance on rotation invariance, therefore, a new matching algorithm based on the SURF interest point and DAISY descriptor is proposed. The method performs better in terms of ro-

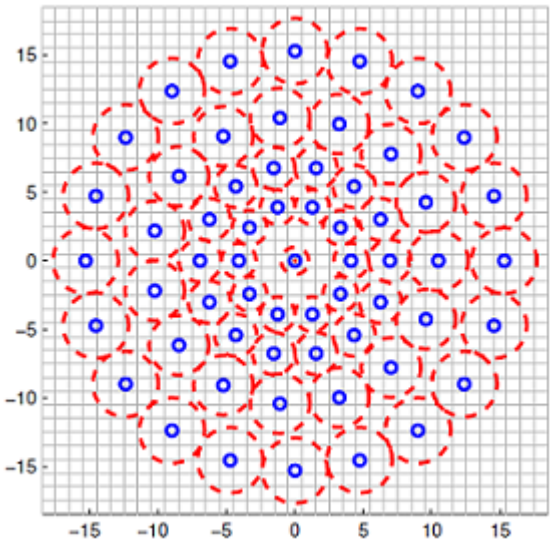


Figure 2.10: BRISK descriptor [7]

tation invariance, but increases the computation time [40]. The reason SIFT and SURF are slow is that these methods need to compute a gradient for each pixel.

To overcome these issues many binary descriptors have been proposed which encode the information into binary strings. In [98], a descriptor named BRIEF (Binary Robust Elementary Feature) is proposed whose performance is comparable to SIFT irrespective of illumination changes, perspective distortion and blur, but is not rotation invariant. To overcome this limitation, the Oriented FAST and Rotated BRIEF (ORB) descriptors were proposed which are similar to BRIEF [99]. They use optimal sampling for selecting the feature pairs. The orientation compensation makes them rotation invariant. Leutenegger et al. [7] proposed a scale and rotation invariant descriptor called Binary Robust Invariant Scalable Keypoints (BRISK). It uses a symmetric pattern in which sampled points are present in concentric circles around the interest points, as shown in the Figure 2.10. Although it is scale and rotation invariant, it requires more computation time compared to BRIEF.

Alahi et al. [100] developed a new binary descriptor named FREAK (Fast Retina Keypoint) which computes faster than SIFT, SURF and BRISK. While

matching the descriptor, the first half bits are compared and if the distance between the first half bits is less than the second half, only then second half bits are compared. This method outperforms ORB and BRISK but not BRIEF in the case of view point change. When scale and rotation change, FREAK slightly outperforms BRISK and BRIEF but behaves worse than ORB in accuracy. FREAK performs worse than the other methods when images are blurred. From the above discussion we can say that Binary descriptors are efficient but they are not affine invariant.

The combination of geometric properties of a feature can be used to make geometrically oriented descriptors. Govind et al. [101] represented the contours extracted from rotated images by slopes of tangents in the contour points. Therefore, for finding correspondence, only proposed descriptions are compared instead of contours and rotation differences between images. In [102], to compensate for rotation difference, angle histograms are computed on line feature points. After eliminating the rotation difference, the feature correspondence is computed by means of correlation. A method to compute the different transformation parameters is developed by combining different features and their descriptors [103]. The geometric deformation between images is divided into elementary steps. Then the transformation parameter value is estimated. Every type of feature descriptor votes for a corresponding value of transformation parameter. The value of the parameter that gets the maximum number of votes is then selected. Another way of finding the correspondence is to compare the Multivalue logical tree (MVLT) formed by using different descriptors of features [104].

A large number of methods have deployed moment-based invariants for descriptions of closed-boundary regions features. Also, affine transform invariants are derived to successfully register landsat and SPOT images [105]. Holm et al. [106] represented the closed boundary regions by their area, perimeter, compactness and moment invariants. Brivio et al. [107] modelled the shadow structures by

means of their inertia ellipses where ellipses are described by their inclination to main axis, area and ellipticity. Li et al. [108] used two Hu's moments for matching closed contours. In [105], the circular moments of the distribution of line feature orientation are used directly to compute the transformation present between the images. For that, scale-space representation and moments are combined together. For registering the SPOT and AVHRR images, a new class of moments invariant to image blurring are proposed in [84].

Zhang et al. [109] have proposed a method in which invariant moments of fingerprints are used for fingerprint matching. Although this method is efficient and rotation, translation and scale invariant, the major problem is that it is based on detection of core points and moments present in the region around the core point. However, for partial fingerprints it is quite possible that they do not have a core point. The quality of the fingerprint may be better in other regions compared to the region around the core point. In addition, in this method common regions in two fingerprints are cropped after aligning the fingerprints. A small piece of information present in the cropped region is used. The rest of information remains unused. Another issue is that this method requires precise detection of the core point in two fingerprints which is very difficult when fingerprints are having intra class variations. The Equal Error Rate (EER) of this technique is 3.64 for FVC2002DB1 (shown in Table 2.1) which is better than other non-minutiae based techniques except the technique proposed in [110]. This is because it has been evaluated on both training and test databases whereas other techniques are evaluated on only the test set that directly influences the performance of the system. Qader et al. [111] also proposed a method based on Zernike moments. This method has faced almost similar shortcomings faced by the method proposed in [112]. In addition, this technique performs worst among all the non-minutiae based techniques with an EER 7.13.

Jain et al. [12] have developed a method which uses the ridge features of fingerprints for alignment and adaptive elastic string for matching the aligned feature set. According to this approach each feature in the fingerprint is associated with the ridge on which it resides. The ridge is represented by a planar curve with an origin at the feature location and the direction of the feature as x-axis. The transformation parameters are computed between the pairs of ridges. This process is iterated until a pair of ridges does not satisfy a threshold. The features residing on matching ridges are used as reference features and used for computing the polar coordinates of the rest of the features. Then, both images are converted into a string of features in the order of increasing radial angle. The strings are compared by using the dynamic programming for computing the edit distance. An adaptive tolerance window is used for tolerating the local distortion. Edit distance between two features tolerates spurious and missing features. Even though an exhaustive reordering and matching is done, that could be because of error in feature extraction.

Feature based methods in a medical field can be further divided into two main categories. (i) Matching among features is established using some criteria based on geometrical properties. Then the transformation parameters are computed based on matching found as shown in Figure 2.11. An example occurs when features are extracted from the input image where each feature is represented by a descriptor. The distance between each descriptor is computed. The similarity score between images is given by sum of all 'corresponding distances' [113] [114] [115]. This approach is reliable when the feature descriptors are invariant to the transformations present between images. (ii) The transformation and matching are based on the optimization of a similarity metric between the extracted features from images. In this case, instead of using pixel intensity for registering the images, the extracted features are used to define the registration.

Category	Method	EER (%)
Fingerprint specific feature based	Kovacs-Vajna, 2000 [116]	4.3
	Tico, 2003 [6]	4
	Chen, 2005 [117]	4.6
	Liu, 2005 [118]	4.3
	Gao, 2011 [82]	3.5
Image feature based	Sha, 2003 [119]	6.23
	Ito, 2005 [110]	1.9 (Private dataset)
	Lumini, 2005 [120]	4.2
	Qader, 2007 [111]	7.13
	Yang, 2008 [109]	3.64
Region based	Karna, 2008 [23]	2 (Private dataset)
	Conventional Correlation [10]	7.1
	Zanganeh, 2014[10] (Average)	2.02

Table 2.1: Performance of the existing fingerprint matching techniques (in terms of EER) on the FVC2002DB1 dataset

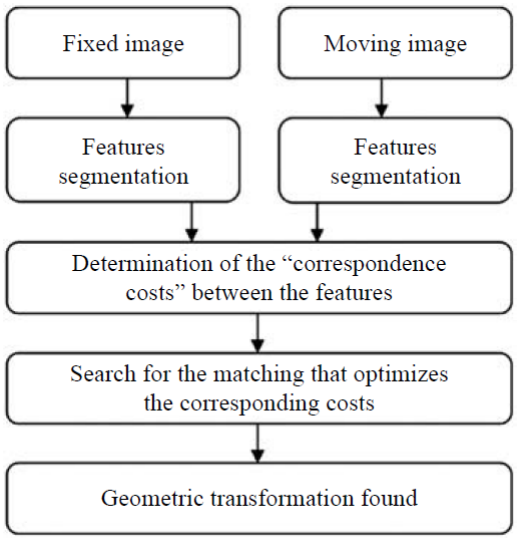


Figure 2.11: Block diagram of the feature based transformation parameter computation approach [8].

2.4.3 Computing Similarity

For computing the similarity between images, the correspondence between features needs to be found. The easiest way to find the correspondence between feature points is to compare all features of one image to all features of another image. However, this approach is quadratic in the expected number of features, which is not practical for some applications. More efficient algorithms can be derived by using the indexing schemes, many of these are based on finding the nearest neighbours in high dimensional spaces. In [121], a method called slicing is proposed that uses a series of binary searches to cut down a number of candidates to compare. Another useful method is the Best-Bin-First algorithm in which a modified search ordering is used for a k-d tree algorithm so that bins in the feature space are searched in the order of their closest distance from query. More recently, an algorithm called Metric Tree has been proposed that compares the descriptors of the features with a smaller number of descriptors at each hierarchy [122].

Once an initial set of feature correspondence is detected, we need to detect a set that will produce high alignment accuracy. One possible method is simply to compute the least square estimate. However, in many cases, it is better to use random sample consensus (RANSAC) [123] and the least median of squares [124].

Once the correspondence between the features is computed, the similarity score is computed for matching images. The similarity score is calculated between sensed and reference images to decide whether both images are matching.

$$Similarity\ Score = \frac{k}{(r+s)/2} \tag{2.11}$$

where k is the number of feature points matched in two images and $(r+s)/2$ represents average number of features in reference and sensed images. While matching it is expected that each feature in an image should have only one matching feature

in the other image. The alignment of images before computing the similarity score reduces the complexity of the matching algorithm.

2.4.4 Summary of the Feature Based Techniques

Feature based techniques are the most reliable techniques for high quality images [9] [3]. It is shown that feature based techniques are highly accurate for high quality images, but these techniques perform badly in the case of low quality images [125]. This is because, for low quality images reliable detection of features is very difficult. Moreover, for low quality images these methods suffer from detection of spurious features due to unclear details in images. Although, the problem of false features can be solved by detecting the features only from valid and good quality regions, the problem again arises if the image does not have any valid regions.

The information used in feature based techniques is quite limited because they use properties of only some points in an image and information in the vicinity of these points. The use of limited information affects the accuracy when a system has to deal with partial images as the partial images already have lost some significant information. Feature based techniques even use some of the remaining available information. In partial images, there exist few features and that is why feature based techniques perform poorly [3]. Moreover, a few features are not enough for distinguishing an image from the inter-class images in the applications of image recognition.

The feature based methods for partial image matching have also been developed. In [126], a Support Vector Machine is used for feature-based partial fingerprint matching. Jea et al. [127] have used secondary features for matching the partial images. Secondary features have been extracted by using the two nearest neighbours' information. However, the error rate of these techniques is quite high and these methods are only evaluated on the FVC2002DB1 dataset, whereas other

datasets of FVC2002 consist of distorted images. Therefore evaluation on these datasets is required as well.

Global matching of images using the features of images are computationally expensive and suffer from the presence of non-linear distortion. Whereas, local feature matching methods depend on neighbouring features and use the attributes which are invariant to linear translation. However, they suffer from lower discriminative capability compared to global matching methods. Both global and local techniques are good in their own way but have limitations as well. Therefore, implementing hybrid strategies can have a profound effect on the accuracy [5].

2.5 Pixel based Techniques

Instead of using specific features of images for matching, pixel or textural based information of images can be used for matching. These methods are intutive and take full advantage of given pixel information [40]. The methods which use the pixel information to match the images are generally called direct methods or correlation-like methods, as opposed to feature based methods [9]. For image alignment, the basic operations are extracting the features, matching the features and finding the correspondence between images. In these techniques, no features are extracted as pixel intensity is compared directly, hence the feature extraction step is combined with the feature matching step.

The global and local pixel/textural information present in images is a good alternative to features. For example, ridge textures present in fingerprints can be used to match them. The ridges in fingerprints are represented by orientation and frequency except at singular regions. Singularities and minutiae are the discontinuities in fingerprint textures. The global information combines the contributions of different local regions into a global measurement and therefore, loses the spatial information. However, local information is mainly related to the detailed informa-

tion of an image which is generally extracted by the use of a specialized bank of filters. For poor quality images, direct methods perform well compared to feature based methods because pixel based information is less affected by the quality of an image compared to specific features [128].

To match the images using direct methods, an appropriate similarity metric needs to be chosen first to compare images. Many similarity metrics have been developed for different matching applications and some of them are discussed below.

2.5.1 Similarity Metrics

The most commonly used similarity metrics are based on intensity cross-correlation, intensity difference or information theory. One of the commonly used and easiest similarity metrics is the correlation between images. Correlation establishes pixel correspondence between two images by shifting one image with respect to the other. This metric of similarity is computed between windows of reference and sensed images. Then, the window of the images that results in a maximum correlation value is searched. The window pair that results in maximum similarity value is taken as a reference window. A correlation based method can align images exactly only if the transformation between images is translation. However, the correlation is a computationally exhaustive process and the complexity increases with an increase in transformation complexity.

Correlation based similarity can be calculated in many ways, such as the sum of absolute differences, sum of squared differences, sum of Hamming distance, normalised cross correlation (NCC) and so on. Let us suppose I_R and I_S are reference and sensed images, respectively. The diversity between two images can be measured as sum of squared difference (SSD) between intensity of images [3]:

$$SSD(I_R, I_S) = ||I_R - I_S||^2 = ||I_R||^2 + ||I_S||^2 - 2I_R^T I_S \tag{2.12}$$

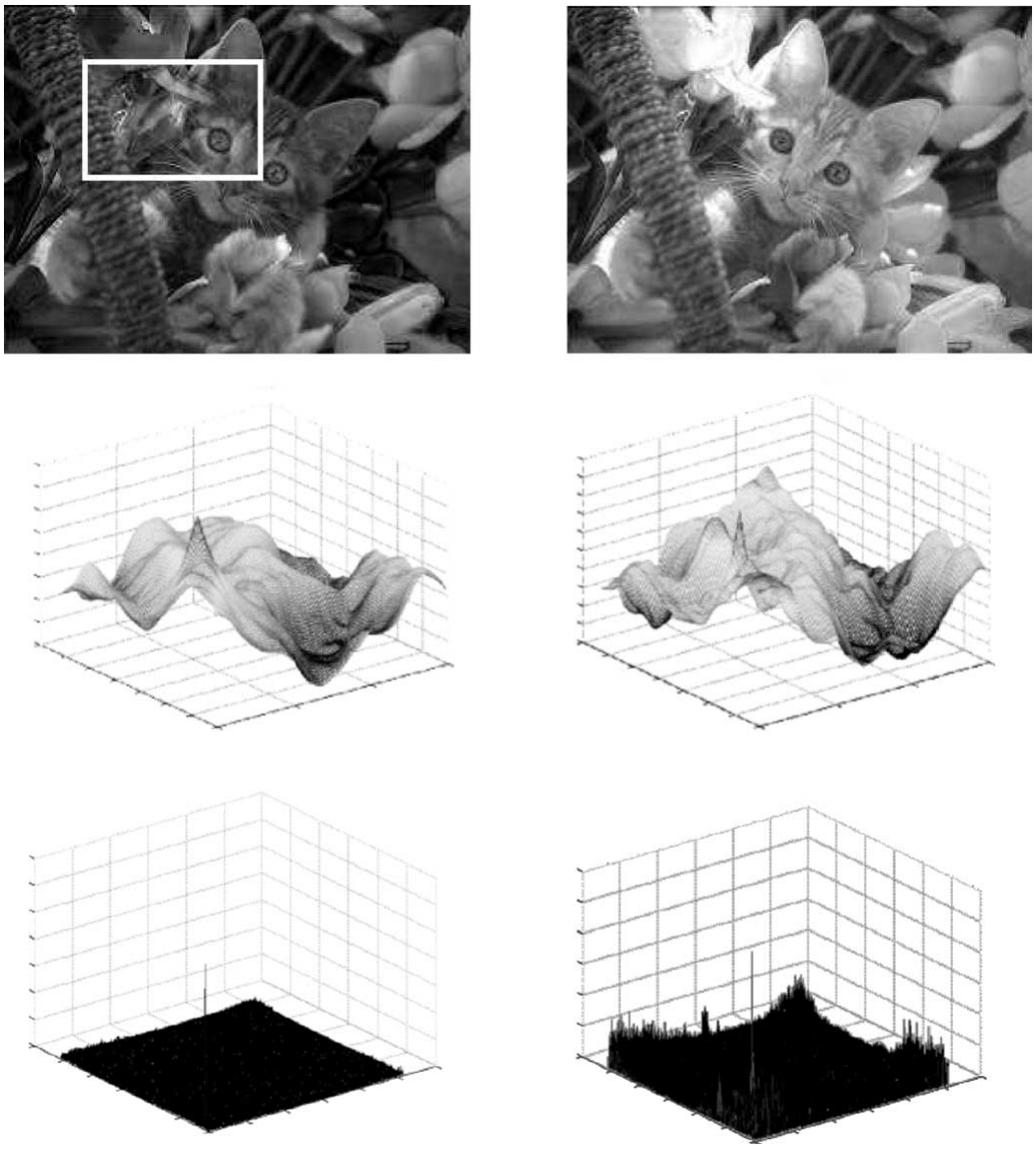


Figure 2.12: Pixel based matching methods: small windows of images are aligned by using normalised cross correlation (middle row) and phase correlation (last row). The global maximum identified in both graphs is the matching location. [9]

where the intensities of images I_R and I_S are represented as vectors of intensities of pixels and T represents transpose. If the components $||I_R||^2$ and $||I_S||^2$ are constant then the diversity between images is minimised when the cross correlation (CC) between two images is maximised:

$$CC(I_R, I_S) = I_R^T I_S \tag{2.13}$$

The sensed image suffers from displacement and angle differences with respect to the reference image. Therefore, correlation between two images cannot be simply calculated by superimposing two images. Let $I_S^{\delta x, \delta y, \theta}$ represent a rotated and displaced sensed image, where δx is displacement in the x direction, δy is displacement in the y direction and θ is the rotational angle. Then the correlation between two images is calculated as

$$CC(I_R, I_S) = \max_{\triangle x, \triangle y, \theta} CC(I_R, I_S^{\delta x, \delta y, \theta}) \tag{2.14}$$

The intensity of pixels can affect the correlation value, hence, it is necessary to adjust the intensities of every image according to the desired mean and variance [10] (example shown in Figure 2.12). In other words, images need to be normalised to calculate correlation independent of intensities. The normalised cross correlation (NCC) is more complex [10] [129] and is given by:

$$NCC = \sum_{x,y} \frac{(I_S(x,y) - \bar{I}_S)(I_R(x,y) - \bar{I}_R)}{\delta_{I_S} \times \delta_{I_R}} \tag{2.15}$$

where $\bar{I}_S$ and $\bar{I}_R$ represent the mean intensities of sensed and reference images and δ_{I_S} and δ_{I_R} are the standard deviations of sensed and reference images, respectively. The correlation value is used as a similarity metric between two images. If the similarity between two images is higher than a threshold, then the sensed image will be accepted as a matching image. Otherwise, it will be rejected.

Correlation mainly has two limitations: high computational complexity and flatness of similarity metric maxima. For sharpening the maxima, different preprocessing techniques can be applied to images. However this will increase the computational complexity of the matching process further. In spite of these limitations of correlation, it has been widely used in different applications because the simple hardware implementations of correlation are useful for real time applications.

As the correlation process is time consuming, other metrics have been developed. One of such metric is the sum of absolute difference (SAD) which is generally used in motion estimation for video coding and is given by:

$$SAD(I_R, I_S) = \sum_{x,y} ||I_R(x,y) - I_S(x,y)|| \tag{2.16}$$

However, this function is not differentiable at the origin so cannot be used for gradient descent approaches. Therefore, a function is required that varies smoothly and is quadratic for smaller values and grows slowly from the origin. One of such function is discussed in [130] called the Geman-McClure function, which is given as

$$\sigma_{GM}(x) = \frac{x^2}{1 + \frac{x^2}{a^2}} \tag{2.17}$$

where a is a constant that can be considered as an outlier threshold. Robust statistics such as the median of absolute differences can be used for computing the appropriate threshold value for this method [131].

The similarity metrics discussed above ignore the fact that some pixels in an image have more importance than others. Therefore it is required to partially or completely downweight the contribution of some pixels [9]. For example, in mosaicing applications the undesired foreground objects need to be cut out. For that, it might be required to erase some parts of images. Similarly, in background

stabilization, we need to downweight the middle part as the middle part consists of moving objects captured by the camera. This is achievable by applying varying weight to pixels of both images to be matched. Then the error between the two images can be computed by using the weighted SSD function given by:

$$WSSD = \sum_{x,y} ||W_R I_R(x,y) - W_S I_S(x,y)||^2 \tag{2.18}$$

where W_R and W_S represent the pixel weights in reference and sensed images, respectively.

Similar to correlation methods is the sequential similarity detection algorithm (SSDA), which is based on sequential search and a computationally simpler distance metric rather than correlation [132]. It sums the SAD of a window pair and uses a threshold criteria. If the sum value exceeds a threshold value then the pair is rejected and the next pair is tested. This method is less accurate compared to correlation but faster [9]. Irrespective of the normal correlation, correlation ratio based methods have been developed in the area of multi-modal registration. This similarity metric can handle the intensity differences between images due to usage of different sensors (multi-modal image). In the case of projection-based registration in fixed-pattern images, this similarity metric outperforms the standard correlation.

The metrics based on pixel intensity difference are SSD or the SSD normalisation [133][134]. The assumption for using SSD is that both images should have identical intensities. The correlations and their variants have also been used as a similarity metric for computing the transformation parameters between images [135][136]. Moreover, cross correlation is based on the assumption that there is a linear relation between the intensities of the images. Hence, the larger the cross correlation between images, the better the reference image. The correlation, cross-

correlation and SSD are the appropriate similarity metrics for mono-modal image registration.

As aforementioned, depending on the structures extracted from original images, the pixel intensity based similarity metrics can be used for image registration. For example, in medical applications, after segmenting the object from images, the pixel based intensity can be used for registering the images instead of using the shape of object [137]. Similarly, when a segmentation process divides the image into smaller patches, the distance or similarity between these patches is computed by using intensity based similarity metrics. In [138], the correlation ratios are considered as the similarity metrics, to register the sets of fibres extracted from brain white matter images.

A variant of regular correlation called phase correlation is discussed in [139]. The phase correlation outputs a single impulse at the correct displacement (as shown in Figure 2.12) and makes it easier to estimate the correct displacement. It performs better compared to regular correlation but the result depends on the noise present in the image.

In [140], another registration method based on Hanusdorff Distance (HD) is proposed to register the binary images transformed by translation or rotation plus translation. The results are compared with correlation results and results show that HD performs better for images with perturbed pixel locations as compared to correlation.

The information theory based similarity metric is Mutual information (MI). It is based on the Shannon entropy and is computed from the joint probability distribution of image pixel intensities. It measures how well one image explains the other image and it can be applied for registering both intra and inter-model images. MI has proved to be a very robust and reliable similarity metric. However, it faces problem for registering small sized images. To overcome this limitation, a method is proposed in which MI is used for global registrations and cross-correlation is

used for registering the small image patches [141]. In multi-modal applications, intensity based registration is used. MI as a similarity metric plays an important role in this process [142].

Another major pixel based image matching method is matching images by computing geometric transformation parameters of images [143]. The matching can be established iteratively either by using the transformation parameters computed or without using parameters [144]. However, in both situations, a similarity metric needs to be optimized. For iterative matching of the images, the common algorithms are Iterative Closest Point (ICP) and its variants [145][146]. In these algorithms the matching is done without using transformation parameters and are based on optimization of a distance or similarity metric. The distance between the features to be matched is based on their particular characteristics. Dedicated optimization methods can be used for establishing matching such as simulated annealing [147], self organising map [148], dynamic programming [149].

2.5.2 Search Methods

Once a metric is finalised, a suitable search technique is required. Finding maximum similarity or minimum dissimilarity is a multidimensional optimization problem where the degree of freedom of expected geometrical transformations represents the number of dimensions. The simplest and only method resulting in a global extreme solution is full-search or a brute-force search in which all alignments are tried exhaustively. This approach is commonly used in block matching in motion compensated video compression where a large number of possible options are tried [9]. Although it is a computationally exhaustive approach, it results in highly accurate results.

In practice, this approach is slow, therefore coarse to fine hierarchical search approaches have been developed. For these approaches the results for lower levels

can be used for a finer search at the next higher level [42]. However, this search does not guarantee the same results as produced by full search [9].

In case of images with translation difference, different optimization techniques are developed to find the maxima. However, the limitation of these methods lie in their basic idea [9]. Generally a rectangular window is used which can easily register the images which only differ in translation. However, if some complex transformations exist between images then this method cannot register the images. Alternatively, different transforms or incremental methods based on the Taylor series expansion are also used for matching the images [150].

2.5.3 Types of Pixel Based Matching Techniques

Different methods have been developed in literature which use the pixel information of images for matching the images. These methods can be broadly classified into three categories namely, correlation based, transform based and mutual information based. The details of these methods are given below:

2.5.3.1 Correalation Based Methods

In these methods, images are compared by computing the correlation value between them. Two images are superimposed and the correlation between images is calculated for different displacements and rotations. These techniques use grey scale information directly from images. The correlation coefficient lies between -1 to $+1$: a value closer to $+1$ quantifies that the images are highly correlated. Generally, images are aligned to make them rotation and displacement invariant before computing correlation [151]. Moreover, it is vital to normalise the grey scale image as images may have been captured under different illumination conditions.

However, the direct application of Equation 2.14 and 2.15 for calculating the similarity between two images is computationally very expensive because of a large

number of mathematical operations [3]. Methods such as the Fourier-Mellin transform provide rotation and translation invariance that can reduce the computational complexity but at the same time affect the accuracy [152]. Also, Fourier transforms or fast optical correlators have been used for improving the speed. However, these techniques do not improve the performance significantly [3].

Computational complexity can be reduced by aligning the images before computing the correlation between them [3]. One such method is developed by Karna et al. [23] which is similar to the method developed in [110]. The only difference is that instead of calculating the phase correlation, normalised cross correlation is calculated between the common regions of the two images. The EER for this method is smaller than the feature based techniques [3]. However, there are several limitations related to this method. Firstly, in this method, images are aligned manually by selecting a common point in both sensed and reference images. Although this helps in aligning the images accurately, it is not feasible in real time applications. Secondly, extracting the same region accurately is not a trivial task. If the extracted regions are not the same it will have direct effect on the matching accuracy. Moreover, even if the extracted regions are similar enough, the method is not able to tolerate the non linear distortion and brightness [151], because similarity is measured globally. The effects of distortion get incorporated while computing correlation. That deteriorates the correlation value. Moreover, matching the partial images by global correlation based techniques is still an unresolved issue. Also, for correlation techniques which require pre-alignment of images, the accurate detection of a feature point is quite challenging.

Zanganeh et al. [10] have developed a region based fingerprint identification method that is robust and accurate. This approach basically consists of three steps: fingerprint alignment, common region extraction and similarity computation. The accurate alignment of images is very important in order to have pixel-to-pixel correspondence between images. Even a slight rotation difference between images

can result in a false matching decision. In this approach, fingerprints are aligned either by using feature points or brute force search. In the case of alignment by using image features, partial and low quality images face problems. In partial fingerprints, it is quite possible that the feature point required for aligning the images is not present. In bad quality images, it is very difficult to detect the feature points accurately. Next, when fingerprints do not have feature points then fingerprints are aligned by an exhaustive search that is very time consuming. For that, the sensed image is rotated for angles between $-\alpha$ to α degrees to find a correct angle for aligning images. The value of α can be 35 or 15 degrees depending on reference points being present or not in fingerprints. The value of α depends on the database as well. The alignment process starts with cropping a region from a sensed image. The size of the region depends on the valid fingerprint region. A sliding window approach is used for template fingerprints in which region around each pixel is considered. Then, for each angle between $-\alpha$ to α the following steps are performed to find the angle at which these two fingerprints can be aligned.

- Rotate the sensed image region by the angle α.
- Compute the correlation coefficients between the sensed image region and all the regions of reference image.
- Select a reference image region as the common region that results in the maximum correlation value.
- Select the rotation angle that gives the maximum correlation value. Then align the images rotationally by using this angle.

The algorithm to align fingerprints by this approach is shown in Figure 2.13 [10].

The common region is considered as a common feature of fingerprints. The size of the common region can vary depending on the sizes of valid fingerprints.

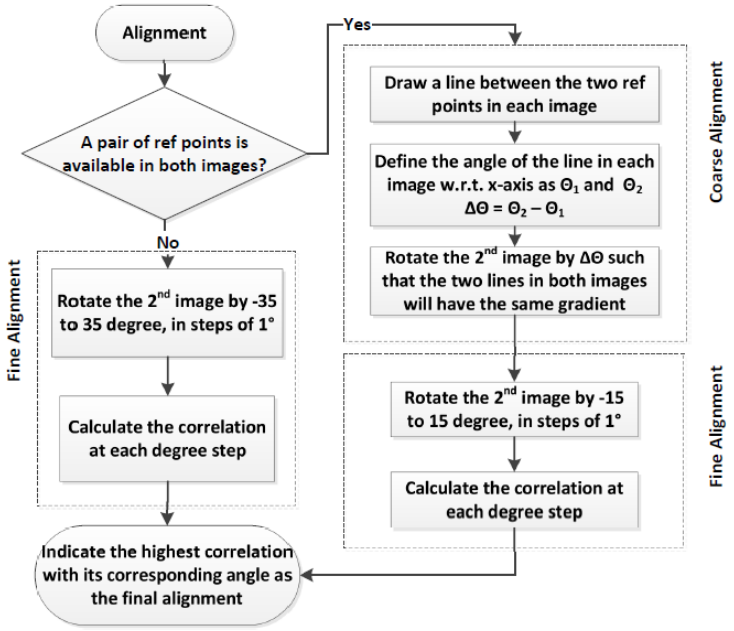


Figure 2.13: Block diagram of the image alignment method with or without using image features [10]

Also, regions can be cropped from different parts of the fingerprint. These properties make this strategy applicable to partial or low quality fingerprints. Once the fingerprints are aligned, the images are divided into the regions by using the reference point. Common regions are found by superimposing two images. After detecting the common regions, a similarity score is calculated between the regions. First, local similarity is measured between the regions. This is followed by global similarity calculation. The global similarity is the average of all local similarities. Local similarity is taken into consideration for minimizing the effect of varying quality within the fingerprint. This technique has improved the equal error rate by three times as compared to single region correlation based technique. The EER of the averaging method reduces to 2.02 as compared to a 7.1 EER for the conventional correlation method for FVC2002DB1 dataset (as shown in the Table 2.1). This technique is able to handle non-linear distortion, as similarities are computed at a local level and consolidated for computing the final similarity score [10]. This

method aligns images accurately, but is computationally expensive. That makes it impossible to use in real time applications.

2.5.3.2 Transform Based Methods

Many approaches based on transforms have been proposed in literature that do not require the extraction of features of images. These approaches compute the underlying transformation parameters using the properties of transforms for aligning images. Recently, the Radon and Fourier transforms have been used by researchers for computing the transformation parameters.

The alignment methods based on Fourier transforms rely on the fact that a Fourier transform of the shifted signals has same magnitude but a linearly varying phase. This property helps to quickly find the rotational difference between images. Another useful property of the Fourier transform is that the convolution of two signals in a space domain is comprised of the multiplications of their respective Fourier transforms in the Fourier domain. Thus, to efficiently compute the error between two images over the full range, we can simply use the convolution property of the Fourier transform. Fourier based methods are preferred over correlation when improvement in computational speed is needed, images are acquired under different conditions or images are corrupted by frequency-based noise.

The phase correlation is based on the Fourier transform which is originally developed to register the images transformed by translation. The cross power spectrum of both unaligned images are computed. Using the peak locations the translation difference between images can be computed (as shown in Figure 2.12). This method is robust against non-uniform or time varying illumination changes and frequency dependant noises. The phase correlation method is also exploited for 3D images [153]. In [154], for down-sampled images analytic expressions of phase correlations are derived. The method performs outstandingly for sub-pixel translation estimation across different spectral bands. However, the closed form

solutions introduce misleading impressions that the methods can give sub-pixel shifts to any accuracy, whereas the accuracy is limited by signal to noise ratios in phase correlation. In [155], extension of the phase correlation is developed to deal with translation and rotational differences between images. If the images are scaled as well, then images can be registered using the combination of phase correlation with polar-log mapping of the spectral magnitude [156] [157]. This method has been used to register the affine distorted images [158].

As discussed earlier, to speed up the search process, a hierarchical approach can be used. However, in cases when the search range is equivalent to a significant portion of a big image, the hierarchical approach does not work well because it is generally not possible to coarsen the representation before significant features of an image get blurred away. For such cases, Fourier transform based approaches can be used. The Fast Fourier transform can compute the transform of images (size $(N \times M)$)) in $O(NMlog(NM)$ operations which is significantly faster than the number of operations required for a full search that is $O(N^2M^2)$ [159].

The Fourier transform can be used to accelerate the computation of correlation and sum of squared distances [9]. However, the Fourier transform is applicable only when the correlation or SSD is performed on full images having pixel-to-pixel correspondence. While this is acceptable for small shifts or small sized images, it does not make any sense when the overlap between two images is small or one image is a subset of another image. In this case, the cross correlation between two images needs to be replaced by windowed cross correlation.

The Fourier transform is generally used for computing translations between images. However it can be used to compute the rotational difference between images by converting the images into polar coordinates. De Castro et al. [155] have proposed a method to align the images using the finite Fourier transform. However, the method is only applicable when images have large overlapping regions which indicates a small translational difference.

The methods based on the Fourier transform can compute the translational alignment to the nearest pixel, however for different applications such as image stabilization, image stitching etc. higher accuracy is demanded. To obtain the subpixel accuracy, it is possible to estimate different discrete values of location around the best value found and interplote to find the best result [9]. One of the commonly used approaches is to compute gradient descent on SSD by using a Taylor Series expansion of image function [35]. However, the effectiveness of the approach is affected by the quality of the Taylor series approximation. When it starts far away from true displacement, many iterations are required; when it starts at a nearby location to the true location, then few iterations are required which indirectly influences the computation time required.

The cross-correlation and SSD based similarity metric can be evaluated in the frequency domain efficiently using the Fourier transform and its properties. Both these metrics can be evaluated using a shift function. This approach is less time demanding than the iterative optimization technique [160] [161]. Moreover, the scaling and rotation of images can also be achieved by transforming the image into polar coordinates systems. Well known phase correlation can be used for registration of the images [161].

The Hough Transform [75] is also a popular method for feature matching. It converts the problem of point matching to a problem of detecting the highest peak in the Hough space of transformation parameters. Fourier transforms and Wavelet transforms have been used for registering the images [162].

Recently, the Radon transform has been used for estimating the RST (Rotation, Scale and Translation) parameters [163]. Estimation of the transformation parameters linking to images using a Radon projection has gained attention recently. Actually working directly with Radon projections eliminates the need for reconstruction of the images which is prone to reconstruction artefacts and is computationally exhaustive. Chen et al. [19] used the Radon transform with the

Fourier transform to register the images. A Fourier transform is used to find the estimate of spatial shift between images; a Radon transform is used to find the rotational shift between images. A property of a Radon transform is that sinograms of two rotationally unaligned images are similar except for the circular shift along the direction. The correlation between the rows of a Radon transform matrix is used to estimate the rotation angle between images. The results of this approach are less accurate for some images [39]. In [37], the registration method, using a Radon transform is developed for non-destructive testing, radiography images and computed tomography projections. Mooser et al. [38] have developed a method that can deal with all degrees of freedom of transformation. Parameters are estimated as an optimization problem which uses a region based optimizer. However, the error in estimation of both parallel and fan beam geometries is very large. Hjouj et al. [39] developed a technique to estimate reflection, scale, translation and rotation for binary images by using the Radon projections. However, the issue with this technique is that the objective function used for computing the angle cannot be generalized for any angle between 0 to 2π. In [164], the results of this method are used to identify the parameters for fast matching images, however the rotation angle computed is less accurate than using the method in [39].

In [165], another method is proposed to estimate the parameters using the Radon projections. Later in 2015, Nacereddine et al. [14] have proposed a modified version of [165] for estimating the RST parameters relating two images only by using the Radon projections. The authors have defined the 2π Radon Transform to deal with the issue of a rotation angle lying between $[0, 2\pi]$, unlike in [39]. This method finds the underlying transformation parameters between two images, and gives the best similarity between the two images by estimating the geometric parameters. However, the method of computing the rotational angle is time-consuming with a computational complexity of $O(P^2N)$ where P represents

the number of projections between 0 to 2π and $N \times N$ represents the size of an image that slows down the entire registration procedure.

In [11], a rotation invariant texture classification approach based on Radon and Wavelet transforms has been developed. This approach takes advantage of the fact that texture patterns do not have any specific direction. The Wavelet energy features are variant with direction, therefore, texture need to be aligned rotationally. For aligning textures, the principal direction of textures is found by using a Radon transform. However, before the Radon transform can be used, a disc shaped area is cropped from the image to make the method isotropic. The principal direction is defined as the direction in which there exist the greatest number of straight lines. The variance of the Radon transform along this direction has maximum variance (as shown in Figure 2.14). After computing the principal direction the texture is rotated to make the principal direction zero. A Wavelet transform is applied to rotated images to extract the features. The k-NN classifier is used to classify the textures based on the features extracted. A rotation and scale invariant method for texture analysis has also been developed in [166]. Both of these approaches are robust to additive noise and find the accurate direction for anisotropic textures. However, they face a problem when the texture has multiple directions.

In [167] a rotation invariant efficient hierarchical texture analysis method is developed based on the Radon transform and Gabor filters. The directional information is computed in high-band texture images that are extracted by Gabor wavelets. This directional information is used for making the extracted features rotation invariant. However, the parameters of Gabor filters need to be tuned very finely that is very difficult. Cui et al. [168] proposed a method for a challenging problem of classification of rotated and scaled textures. The rotation and scale invariant features are extracted by computing the Radon transform and then applying a transformation invariant adaptive 1-D Wavelet transform in the Radon

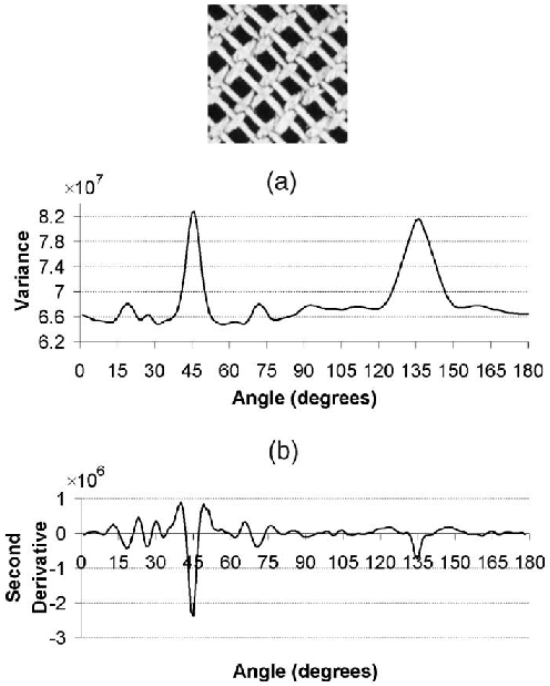


Figure 2.14: (a) Directional texture rotated at 45°; (b) Variance of Radon projections at different angles; (c) Second derivative of (b). The local maximum exists at 45° [11].

domain. The Radon transform shifts the row in the feature matrix whenever the image rotates, so a row shift invariant Wavelet transform is also applied. The feature image is divided into different scales, to make the classification faster. It is compared with different wavelet based methods under different SNRs. It is more robust to noise than other methods. However, this process is computationally expensive.

Xiao et al. [169] have developed a rotation invariant texture analysis method base on the Radon and Fourier transforms. The Radon transform converts the rotation difference in the image to translation, and then uses the translation property of the Fourier transform that states that translation in the original domain converts into phase in the Fourier domain. Therefore, two rotationally misaligned images in the Radon domain have same magnitude but differ in phase. Although, the Wavelet transform is more popular than the Fourier transform, the Fourier transform can be computed much faster and more conveniently. The method is

robust to white noise due to the process of summing the pixels for computing the projections in the Radon transform.

The Radon transform has also been used for handwritten character recognition [170]. The selection of features for characters is very tedious. It becomes more difficult when the characters are rotated. The main objective of this approach is to construct a transformation invariant vector by using the Radon transform and Principal Component Analysis (PCA). The Radon transform transforms the image from the space domain into the Radon domain. Generally, the size of the Radon transform matrix is large. Therefore, the size of the output matrix is reduced by using a resizing operation. Then PCA is applied to compute the rotation invariant vector. The number of local maxima in the Radon transform output are also stored with the vector for the preliminary classification.This method has improved the accuracy as compared to the Optical Character Recognition system.

Most conventional approaches for fingerprint recognition are matching based on fingerprint features. However, the use of the Radon transform for fingerprint recognition is quite a different approach. The rotational property of the Radon transform makes it quite attractive for fingerprint recognition, because one of the major causes of intra class variations in fingerprints is rotational difference. The use of the Radon transform in fingerprint recognition makes it easier to deal with the rotational differences of fingerprints. Some of the Radon transform based fingerprint recognition techniques are discussed in following section.

Sandhan et al. [171], have developed a method based on abstracted Radon profiles which extracts the global properties of fingerprint and does not need extensive image preprocessing. In this approach, each Radon projection is fitted to a polynomial and few coefficients of each Radon projection are captured. The coefficients for all projections are combined together to form a vector for a fingerprint. The vectors for two rotationally misaligned fingerprints are same with a circular shift of the coefficients. The circular shift depends on the angular difference between

two fingerprints. Hence, the test image is analysed for all possible orientations for alignment. This approach provides a more computationally efficient way than the brute-force search alignment approach. Also, the local characteristics are preserved by use of a multi-layer architecture. However, for a multi-layer architecture a core point is used as reference. That is difficult to extract. Also, all the fingerprints may not have core points, therefore this method is not applicable to all types of fingerprints.

2.5.3.3 Mutual Information Based Methods

For multimodal image matching, Mutual Information (MI) based methods are one of the leading methods [9]. Registration of multi-modal images are a difficult task. However, they are necessary especially in medical image registration. Applications of image registration in medical fields include fusion of images obtained from computed tomography (CT), magnetic resonance imaging (MRI) with functional images from single-photon emission computed tomography (SPECT) and positron emission tomography. Image registration is required for almost every anatomical part of the body [8]. The comparison of the functional and anatomical images of a patient can result in diagnosis, which otherwise is impossible to get. Moreover, this type of registration is required in remote sensing as well. The Mutual Information is the statistical dependency between two datasets and is suitable for registration of multi-modal images. The mutual information between two random variables X and Y can be given as:

$$MI(X,Y) = H(Y) - H(Y/X) = H(X) + H(Y) - H(X,Y) \tag{2.19}$$

where $H(X) = -E_X(log(P(X)))$ is the entropy of a random variable and P(X) is the probability distribution of X. These methods are based on maximizing the MI

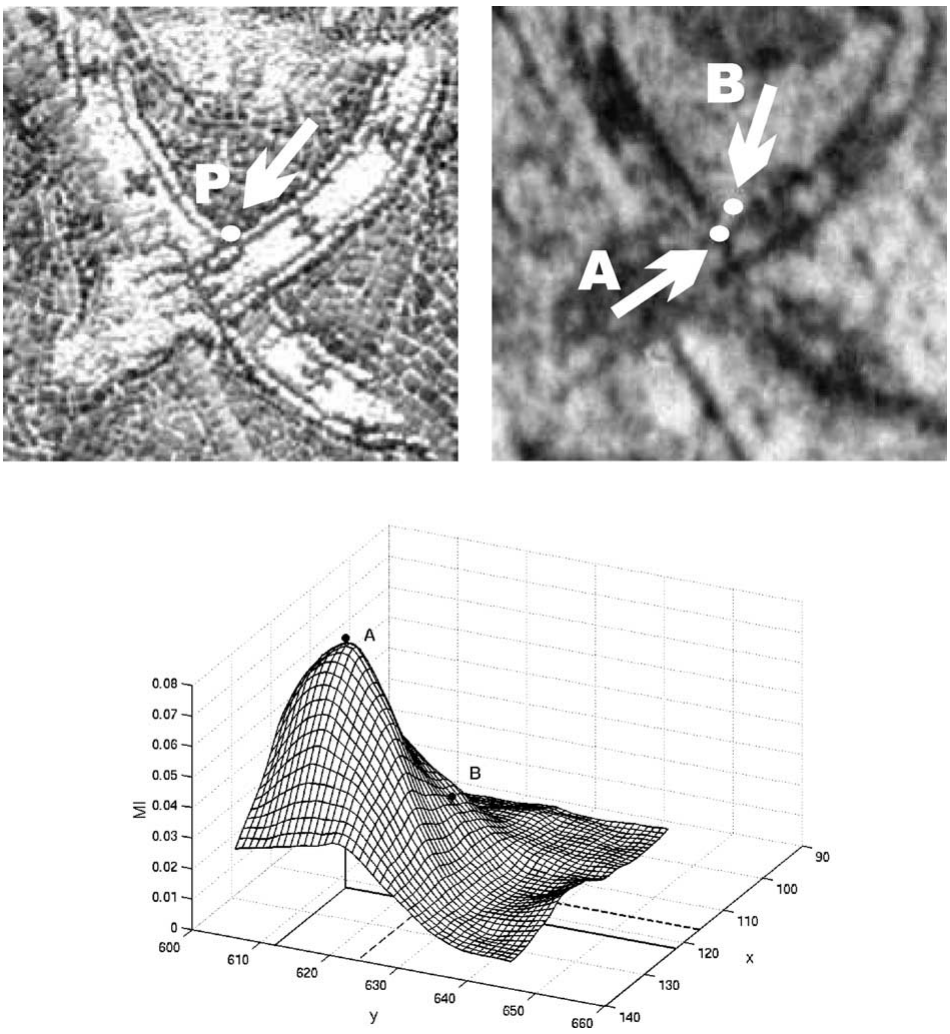


Figure 2.15: The old and new photographs of a mosaic (top row) and mutual information criteria (bottom row). The mutual information gets maximised at point A instead of point B which is selected manually. The mistake is made due to the poor quality of the images [9].

(as shown in Figure 2.15). For improving the efficiency of registration, a coarse to fine approach can be used.

Mutual information based methods are first used for registering the magnetic resonance images and 3D object model matching to real scenes [172]. Later, a hierarchical search strategy with simulated annealing is used to find maximum of MI [173], whereas in [174] maximization is achieved by using a multiresolution hill climbing algorithm and registered MR-CT and MR-PET images. In [175] MI is

compared with other similarity metrics for registering the medical images, however, MI performs worst of all the similarity metrics.

2.5.4 Summary of Pixel Based Techniques

The region based techniques use gray scale images and take into consideration all dimensional attributes of images including both macro and micro characteristics. Therefore, this has the ability to extract all discriminative information present in images as compared to other methods. These techniques do not suffer from false feature detection or shifts in the location of features and hence, have better accuracy than feature based techniques for poor quality images [10]. These techniques are capable of matching the low quality images and have high tolerance to image noise as well. It is also shown that the accuracy can be increased by using the complete information of an image instead of using specific features from that image. The region based techniques can be used for different kinds of images irrespective of images being full or partial (meaning images have or do not have pixel-to-pixel correspondence).

However, there are issues with pixel based techniques that need to be addressed. Most of these techniques need to align the images before matching. For aligning, either a feature point method or an exhaustive method is used. Therefore, a method is required which can either match the images without aligning them or if alignment is required then images can be aligned efficiently without using specific features.

2.6 Hybrid Techniques

In the previous sections, we have discussed the techniques which either use specific features or use pixel based information for matching images. Many techniques

have been developed for matching images by combining both features. The most relevant techniques are discussed below:

Jain et al. [12] have discovered a filter bank based local texture analysis technique. In this technique, an image has been divided into 80 cells (5 bands with 16 cells in each band) around the core point. A vector is formed for each cell representing the local texture information of the cell. An ordered numbering of the cells represents the global information of the image. A bank of 8 Gabor filters is used to obtain the texture information from each cell. Therefore, each fingerprint is represented by 640(80×8) fixed-size vectors which are collectively known as a *FingerCode* (shown in Figure 2.16). The generic term V_{ij} of the vector represents the energy contribution by the filter j in the cell i. The energy is computed as the average absolute deviation from the mean of the response of the filter j over the pixels of cell i.

Two fingerprints are matched by computing the Euclidean distances between corresponding vectors in the two *FingerCodes*. The complete process is shown in Figure 2.16. The disadvantage of this method is that the whole process is focused on the core point. Therefore, if the core point is not detected reliably or it is present near the border, then the *FingerCodes* either will be incomplete or will not be compatible with the reference image. *FingerCodes* consist of some complimentary information but they are not as distinctive as minutiae. However, they can be used with minutiae for improving accuracy. In [176] and [177], this method has been enhanced in which two fingerprints to be matched are first aligned using minutiae and then fingerprints are divided into cells based on the squared mesh grid.

Kovacs-vajna [116], Beleznai et al. [178] and Nandakumar et al. [179] have used pixel information around features.In [110], a phase component based method is developed for fingerprint matching in which a reference fingerprint is compared with the sensed fingerprint rotated at different angles. Before calculating the sim-

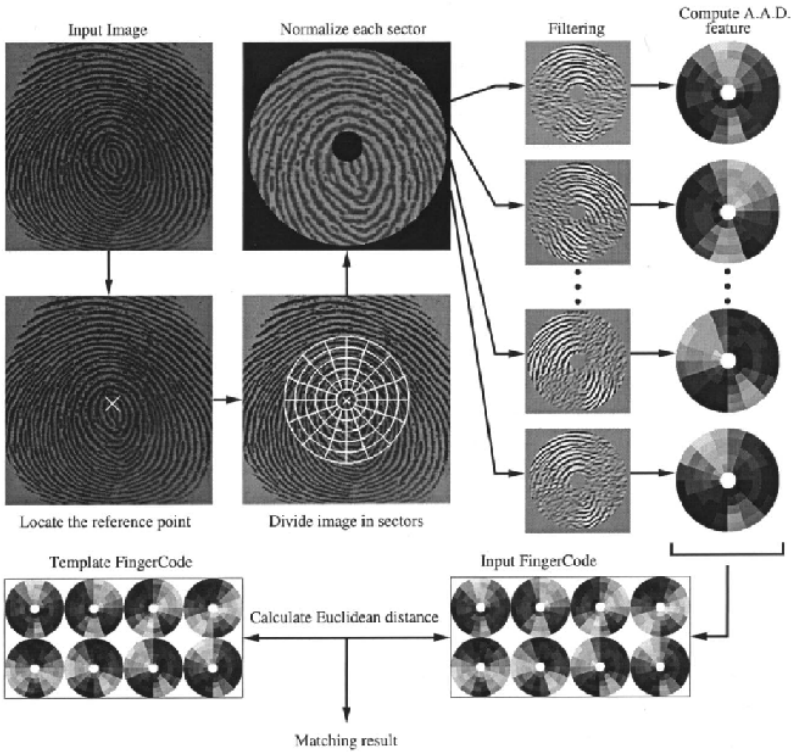


Figure 2.16: System diagram of fingerprint authentication by fingercode [12]

ilarity score, common regions are found in sensed and reference images to improve the accuracy of the system. The EER of this technique for the FVC2002DB1 dataset is 1.9%. Although the error rate of this technique is lower than those of feature based techniques [110], the main issue is that the score is calculated globally. Hence, it is quite possible that it will not work when non-linear distortion is present in image. A hybrid approach based on minutiae and texture information is given in [176]. The results show that fusion of texture and minutiae based algorithms improve the matching score significantly.

Ojala et al. [180] have developed a hybrid technique where images are first aligned by using the features. Then texture information is extracted by Local Binary Patterns (LBP) from the image convolved with Gabor filters. The fingerprint is divided into sub-windows and then the image is segmented to discard the background. Finally, each window is convolved with a Gabor filter and LBP are computed. Two fingerprints are compared by finding the Euclidean distance

between corresponding pairs of sub-windows. Although these techniques capture more distinguishing information from images, they all face the disadvantages faced by feature based and pixel based techniques.

2.7 Learning Based Techniques

The performance of the automated image recognition system is heavily influenced by the accuracy of feature extraction methods [181]. Currently, many techniques have been developed to extract features with acceptable results. Complications start when images are captured in uncontrolled environments that result in low quality and partial images. To deal with these problems, methods based on learning models have been developed. One of the most widely used learning methods is the artificial neural network (ANN). It is inspired from the idea of biological neural networks which have the ability to learn from inputs for making decisions [182]. Artificial neural network based methods are developed for registering images, extracting features, matching images, classifying images and so on.

In [183], an image registration method based on neural networks and Fourier transforms has been presented. In this approach, the spectrum of sensed images are computed and then in a selected window of each spectrum, Fourier coefficients are computed which are fed as a input to neural network. The neural network is implemented to compute the transformation parameters which are used to align images. In [184], another neural network based image registration method is proposed.

Similarly, ANN is also used for logo recognition, however, traditional approaches to logo recogntion use the extracted features and descriptors. One of the recent logo recognition methods based on the convolutional neural network (CNN) is proposed in [185]. The CNN used in this approach is specifically trained for the task of logo recognition. Although the precision-recall of the proposed

approach is comparable to the best approaches proposed in [186] and [187], the process is computationally exhaustive. Teng et al. [188], have proposed a medical image fusion method based on a neuro-fuzzy logic.

Neural networks have been evaluated in fingerprint identification as well. In [189], a new fingerprint matching method based on texture based descriptors and minutiae based descriptors is proposed. For computing the similarity between images, a 17-D vector is computed. A support vector is used to convert a vector into a matching score. The method results in lower EER as compared to existing techniques, however, it faces many challenges. First, the method is dependant on the accuracy of minutiae extraction and hence, in case of partial or low quality images if the minutiae are not extracted accurately, the performance of the method is highly affected. Moreover, the exaction of features and computation of both type of descriptors is a computationally exhaustive process. In addition, for computing the similarity score, the images are required to be aligned and once aligned the minutiae are matched using a greedy approach which increases the computation time of the approach and makes it impossible for real time applications.

Jung et al. [190] proposed a fingerprint classification method that results in an average accuracy of 97.4%. However, before computing the accuracy, a set of 256 fingerprints that are affected by severe noise are excluded in their evaluation. This may mean that, the method might not perform well in case of low quality or noisy images. In [181], a fingerprint identification system based on neural networks is proposed that represents a complete biometric system capable of sensing images, enhancing images, extracting features and matching. The accuracy of feature extraction reaches 98.64% on a mixed dataset of low and high quality images. However, the accuracy depends on the number of training images. Moreover, the overall recognition rate of the method is only 92% and that is even on high quality images.

Although neural networks are used in different applications for improving the performance, there are many limitations as well: (i) the training process of a neural network can be computationally expensive for large datasets that may require high power GPUs; (ii) most existing methods do not incorporate the noisy images (specially latent fingerprints) that constitute the top research issue in forensic applications; (iii) features extracted from images to train the neural network are application dependant, so a neural network cannot be generalised for different image matching applications; (iv) datasets like logo datasets may suffer from the problem from insufficient training datasets; (v) many neural network based techniques depend on image features whose extraction is a difficult task, as partial images may even not have the required features; and (vi) the robustness of ANN based methods is affected by the accuracy of the convergence of training datasets and the quality of images.

2.8 Comparison of Feature based and Pixel based Techniques

In the above sections, we have discussed different image matching techniques. The conventional approaches of image matching can be broadly classified into two categories, namely, feature based and pixel based. These techniques have several advantages and disadvantages that are summarized below in Table 2.2.

Feature based matching methods are (i) used when the information represented by specific points is more significant than the information represented by image intensities, and (ii) recommended only if the images consist of enough distinctive features and easily identifiable objects [191]. These techniques use highly distinguishing local information of images and results in high accuracy in high quality

Techniques	Advantages	Disadvantages
Feature based	• Use highly distinguishing information • Have a high tolerance to non-linear distortion • Are analogous to expert based matching • Do not require Prealignment • Are computationally fast	• Require preprocessing • Have low tolerance to image noise • Suffer from false detection of features • Are not applicable to low quality and partial images • Experience difficulty in accurately detecting features
Pixel Based	• Extract pixel information easily • Are applicable to low quality and partial images • Have high tolerance to noise • Do not suffer from false feature detection • Can use both local and global features • Make optimum use of available information • Are noise tolerant • Do not detect false features	• Use less discriminative information • Have high computation costs • Have low tolerance for non-linear distortion

Table 2.2: Advantages and disadvantages of existing fingerprint matching techniques

images. Moreover, these methods are tolerant to non-linear distortion as features are local information. Non-linear distortion has less effect on the local level.

The extraction of features in images requires image processing techniques such as image segmentation and image enhancement. Therefore, the accuracy of image procesing techniques affect the accuracy of the extracted features' locations [75]. The other major limitation of these methods is that features might be hard to detect or become unstable with time [192]. Moreover, the extracted features need to be discriminative and uniformly distributed in images so that these can distinguish intra and inter class images.

The accuracy of matching depends on the reliability of extracted features and the number of features. The matching accuracy decreases as the number of features decreases. A lower number of features may affect the accuracy. A higher number of features affects the computational time of the matching process. Therefore, if the number of features is high, then post processing is required to select a sufficient subset of features for efficiently matching the images. In addition, some feature based techniques generally do not perform well when images are too textured or not textured enough. In such images, features would often be distributed unevenly and hence fail to match the images accurately [193]. Also, many of these techniques use feature descriptors for matching the images. These feature descriptors are required to be invariant to the assumed difference between images [74]. However, these techniques have a low tolerance to noise. This is because noise can result in detection of spurious features that can affect the accuracy.

The extracted features from the images are used to align them before matching. If images could be pre-aligned then feature matching will be simply a pair matching problem. The features of images can be matched locally or globally. In global matching techniques, images can be aligned by two different ways (i) absolute pre-alignment of the images with respect to a fixed rotation and location (ii) relative pre-alignment of the sensed image with respect to images in database

[3]. Absolute pre-alignment is generally based on feature points which are either difficult to detect or not present in the image. The global matching uses highly distinguishing information from images, however, is time consuming, intolerant to non-linear distortion and has a high computational cost. Therefore, researchers have developed local matching techniques which do not take into account the global transformation like translation and rotation.

In local matching techniques, transformations of images are not considered, because local structures around the feature points that are used for matching are invariant to any kind of transformations. Local structures include pixel intensity around a feature point, angles and distances between features falling in a certain region. No pre-alignment is required in these techniques and hence, these techniques are also called alignment-free techniques. Local matching techniques can be divided into two classes: nearest neighbour based and fixed radius based. Nearest neighbour methods are defined by the relation between selected features and their k nearest neighbours [79]. These techniques suffer from missing or spurious features. In a fixed radius based technique, the neighbours of the central feature are defined by all the features which lie in a circle of radius R to the central feature [83]. Fixed radius techniques are tolerant to missing or spurious features. However, they are more complex and suffer from the border errors. After matching the points based on the local structures, a score is computed based on the consolidation of valid matches to check the validity of local matches at the global level. This consolidation step increases the computation time of the matching process.

The feature based techniques perform poorly in the case of low quality or partial images due to limited use of information. Therefore, specific features alone in poor quality or partial images may not be enough for matching [194]. Maltoni et al. [3] have mentioned the use of additional distinctive features along with the specific features that can increase the robustness and accuracy. When the images have partial or low quality then image pixel based techniques can perform

better than the feature based techniques [3]. Images having more detail have the required number of features for feature based matching. However, images such as medical images are not rich in textures, therefore it is difficult to extract the required features from these images. Hence, for matching medical images, pixel based methods are more suitable [8].

Pixel based techniques are preferred when images do not have enough salient features or precedent details. Detection of salient features in difficult and distinguishing information is provided by gray colors instead of local shape or structures. These methods make optimum use of information present in images as they rely on every image pixel. To remove the effect of noise on pixels, weights are assigned to the pixels. It can be argued that in blurry images with slowly varying gradients, a pixel based approach can find the alignment, whereas feature based methods will fail to align [9].

The use of pixel based registration methods has increased considerably as compared to feature based methods for registering the images. The main reason behind this is the advancement in computational sources, specifically memory capacity and processing speed. Twenty or thirty years ago, computers used to take hours or days to register two images by using intensity based registration. Today, even a simple laptop can be used to solve the registration problem in few seconds by using the intensity based registration method. The other reason for the growing importance of pixel based matching is a consequence of their simplicity as there is no need for image segmentation or other image processing techniques which are complex and prone to errors. Nowadays, because of ease of pixel based methods, these methods are used as an initial step for image registration. If the orientation and position of a structure is known in advance then it is comparatively easy to do matching [8].

Pixel based methods mainly have two limitations: (i) both images to be compared should have same intensity function, either identical or statistically depen-

dant; (ii) existing techniques are able to handle only translation and small rotations between images. Although, it is possible to extend the approach to full rotation and scaling, it is practically of no use due to extreme computational cost. Moreover, specific features represent more distinguishing information than the pixel based information. Therefore, performances of these techniques are worse than that of feature based techniques for high quality images.

The region-based techniques which compose a subfamily of pixel based matching techniques are based on global and local information about images. For methods based on the global information, images need to be aligned first. This is because global features change significantly with the orientation of the images. The images are aligned by using the feature points. Partial images might not have the required feature point or have a feature point whose accurate detection is difficult. This restriction limits the use of this technique in partial images. Also, this method cannot tolerate non-linear distortion as images are matched globally. These issues are solved by region based techniques (such as the method proposed in [10]) that is based on local information. However, it aligns the images by a brute force search method, making it unsuitable for use with real time applications.

2.9 Conclusion

In this chapter we have investigated current techniques for matching the images in different fields including object recognition, biometric recognition, medical image registration and others. The concerns and issues in those techniques are summarised in Table 2.2. From the literature it is apparent that current approaches fail to align and match the images accurately and efficiently.

From the discussion in this chapter, the thesis argues that pixel based matching of images is the most promising technique among the different matching techniques. The use of transforms along with pixel information for matching images simplifies

the matching process. However, the performance of existing pixel based matching techniques, with or without using transforms, still requires improvement. Moreover, as discussed in previous sections, researchers have been working to resolve the issue of image alignment in many image matching applications. However, aligning the images accurately and efficiently is still a challenging issue that affects the overall efficiency and accuracy of the matching process. In addition, most of the transform based methods are proposed for full images which exhibit pixel-to-pixel correspondence and therefore, these methods fail when applied to partial images.

Hence, in response to these limitations, the fundamental question that we want to explore in this thesis is to apply a pixel based approach to match all type of images irrespective of the pixel-to-pixel correspondence between images. Moreover, to eliminate the alignment issue in image matching, we need a method that can match images without explicitly aligning whole images or parts of images. In addition, for applications such as medical image registration where alignment of images is demanded, we need a method that can align images accurately and efficiently.

Chapter 3

Recovery of Similarity Transformation Parameters

3.1 Preamble

Chapter 2, asserts that accurate registration of images is very important [22] for different applications of image processing and computer vision. The applications include medical image diagnosis, object recognition, image retrieval, remote sensing and biometric recognition. Image registration is a process of aligning images into a common coordinate system [25]. This seems a clear problem that should have universal solution, however, this is still away from the state-of-the-art. Due to the wide variety of applications, the registration of images has become a complex problem that has a variety of solutions. Over the past few decades, the research in image registration has substantially increased due to a growing availability of digital images in different computer vision applications.

One of the common methods of registering images is by computing the transformation parameters relating them. The transformation parameters have different uses in different applications such as in object recognition; they help in matching two images while in image registration; in bringing the similar points of registered

images into correspondence. Among all geometrical transformation parameters, similarity transformation parameters namely scale, rotation and translation are often used by different applications where the shape of the object is preserved. Examples of such application include (i) logo dataset (UMD Logos); (ii) pictures taken by camera in a common frame; (iii) weld defect identifications in non destructive testing; (iv) moving object state change quantification; (v) Chinese characters; and (vi) MRI or CT Brain images.

Manual registration is the foremost method of image registration in which the distinctive points are selected in the images manually and then used to compute the geometric transformations between images. This method is widely used in the remote sensing community as it is easy to implement. However, the manual registration is prone to error and laborious, making difficult the processing of large amounts of data [22]. Therefore, researchers have developed automatic methods for computing the parameters. Generally, parameter computation techniques can be divided into two categories, namely, feature based or pixel based.

In feature based methods, the images are pre-processed to extract the distinctive features that are used for aligning images. These methods extract the highly distinguishing information from images and then match them based on local image properties to compute the transformation parameters. Feature based methods can be further classified into two classes: (i) a class of techniques that uses a dense set of low information content features such as edge points, and (ii) a class of features that uses a sparse set of high information content features such as corner points. Numerous researchers have worked on the point pattern matching problem in different fields such as pattern recognition, computational geometry and image processing. One of the preliminary steps in these applications is the detection and extraction of features in images, based on the notion that images can be represented by a set of feature points. Many operators have been developed to detect the feature points (discussed in Chapter 2.4.1). Once the feature points are de-

tected, the local information around the feature points is described by a descriptor for point-to-point image matching and parameter computation [25].

On the other hand, in pixel based methods, the pixel information of regions of images is used to match images. The algorithms either compute correlation between raw pixel values or compute the difference by converting images to other domains such as Radon or Fourier domains [22]. In the absence of reliable features, correlation is a trustworthy method of aligning images. Correlation computes the similarity between images by comparing the pixel intensities. A region from a sensed image is translated to a reference image to find alignment that optimizes the similarity measure. This naive brute-force search approach is used for finding the optimal translation. Moreover, if the images are rotated as well, then the same procedure of finding the similarity is done for different rotation angles as well to find optimal rotation difference between images [10].

Other pixel based methods are transform-based registration methods which work on the premise that information from sensed images makes the geometrical transformations easier to compute. Commonly used transforms are Fourier and Radon transforms. These methods are used for registering images having pixel-to-pixel correspondence. The frequency domain provides a faster alternative for computing the correlation similarity metric. Fourier transform based methods use Fourier's shift theorem that states that if one image is the translated version of another image in the pixel domain, then in the frequency domain they are related by a phase shift which can be computed efficiently. However, this method is only able to find the translation between two images and even in translation it is difficult to compute large displacements. In 1987, DeCastro et al. [155] extended the Fourier method to rotation whereas Chen et al.[156] and Reddy et al. [157] have extended the method to rotation, scale and translation.

Estimation of the transformation parameters linked to images using a Radon projection has gained attention recently. Actually working directly with Radon

projections eliminates the need for reconstruction of images which is prone to reconstruction artefacts and is computationally exhaustive. One of the recent methods based on Radon projections is the method proposed in [14] that estimates the RST (Rotation, Scale and Translation) parameters relating two images only by using the Radon projections and without additional computation. However, the method of computing the rotational angle is time-consuming with a computational complexity of $O(P^2N)$ where $P \times N$ is the size of Radon Matrix (where P is the number of projections and $N \times N$ is the size of image) that slows down the entire registration procedure.

As discussed above, for images having exact pixel-to-pixel correspondence many image registrations methods have been proposed [discussed in Chapter 2]. The feature based methods need to extract the features for computing the parameters, however, accurate feature extraction itself is a difficult task. Moreover, the extraction of features becomes more difficult when images are either partial (images do not have pixel-to-pixel correspondence) or low quality. This is because, partial images might not have the features required for computation of parameters, whereas in low quality images there is a high chance of extracting spurious features that can affect the accuracy of the system. The feature extraction process needs image processing for extracting features accurately which increases the computational complexity of the system. Most importantly, different types of features are required for different applications, hence, a common method cannot be developed for different applications.

Some of these issues are dealt by pixel based techniques as no specific features are required to be extracted. Therefore, the approach can be applied to different applications and extensive image processing is not required. In addition, as no specific features of images are used, there is no chance of false feature detection. However, there are some challenges: brute-force searching is used for computing the transformation parameters which is computationally exhaustive. Furthermore, the

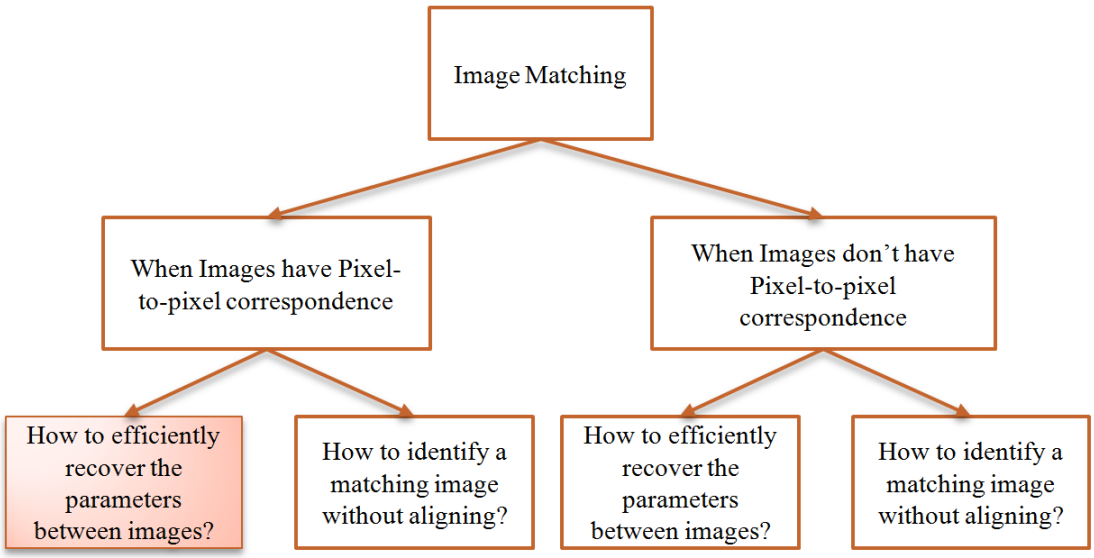


Figure 3.1: The coloured block represents the thesis objective addressed in this chapter.

number of operations required for computing the correlations depends on image sizes, which implies the larger the image the higher the computation time and vice versa. Similarly, transform based methods compute the parameters with high accuracy, however, are computationally expensive.

Although pixel based methods have many deficiencies, the approach is vital and remains widely used in different applications such as remote sensing applications. Moreover, the approach is applicable to low quality images. Therefore, we need a pixel based method for image registration but that should be able to overcome the limitations of computational complexity. For solving this issue, in this chapter we have proposed a parameter based registration method for images having pixel-to-pixel correspondence (shown in Figure 3.1) that can compute the parameters accurately and efficiently with respect to computational time.

The remainder of the chapter is structured as follows: Section 3.2 discusses the parameters to be computed which is followed by the transform used for computing the parameters between two images in Section 3.3. The method of computing the parameters is discussed in Section 3.4. The computation time of the proposed approach is discussed in Section 3.5. The performance of the proposed approach

is tested on images of varying classes and results are presented in Section 3.6. Finally, a conclusion is drawn in Section 3.7.

3.2 Parameters to Compute

An instance of image registration is to align an image with respect to the other image. For this we need to compute the transformation parameters relating to those two images. We can define the image registration problem by three elements:

i) Reference image $I_R(y_R, x_R)$ that is generally taken to be unchanged.

ii) Sensed image $I_S(y_S, x_S)$ which is the transformed version of the reference image.

iii) Geometric transformations that map the spatial positions in the reference image to the sensed image.

The sensed image denoted by $I_S(y_S, x_S)$ is obtained by three functions formulated as:

$$I_S(y_S, x_S) = Translation \times Scale \times Rotation \times I_R(y_R, x_R) \tag{3.1}$$

where (y_S, x_S) are row and column pixel locations of the sensed image I_S and (y_R, x_R) are the corresponding row and column pixel locations in reference image I_R. The coordinates of the sensed image can be represented as a matrix product using homogeneous coordinates:

$$\begin{bmatrix} y_S \\ x_S \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & y_0 \\ 0 & 1 & x_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi_0 & -\sin\phi_0 & 0 \\ \sin\phi_0 & \cos\phi_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_R \\ x_R \\ 1 \end{bmatrix} \tag{3.2}$$

where (x_0, y_0), α and ϕ_0 represent the translation (in R^2), scale (in R_+) and rotation (in $[1, 360]$) parameters, respectively. Hence, the problem here is to devise

an algorithm for all pairs of images that will compute transformations for any instance. The objectives are to (i) speed up the registration process to handle bigger datasets; and (ii) improve robustness and reliability of the registration algorithm so that it can tolerate noise.

3.3 Transform for Computing Parameters

The information represented in the transformed domain makes it simpler to compute the similarity transformation parameters. One such commonly used transform is the Radon transform. The Radon transform converts the image from the pixel domain to the Radon domain and makes it easy to handle the rotational difference between images.

The Radon transform (RT) was discovered in 1917. Since then, it has been deeply studied and applied in different areas, especially in biomedical imaging fields such as reconstruction of images from the projections in computer tomography. The use of the Radon transform in image processing has gained momentum recently. The most recent use of the Radon transform is for computing the underlying transformation parameters relating to two images directly from their projections only, assuming no knowledge of the reference image [14].

The Radon transform $RT_f(\phi, u)$ of a two dimensional real valued function $f(x+jy)$ along a straight line $Z(t) = e^{j\phi}(u+jt)$, where t is the line parameter and given by

$$RT_f(\phi, u) = \int_{-\infty}^{\infty} f(Z(t))dt \tag{3.3}$$

In other words, the Radon transform computes projections of a $2D$ function along defined directions (as shown in Figure 3.2), where a projection represents a set of integrals along a defined direction [195].

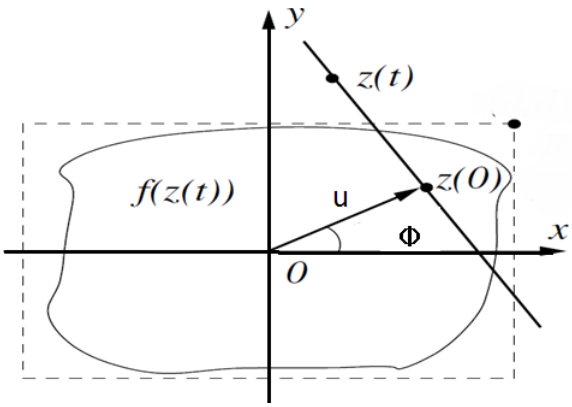


Figure 3.2: The computation of the Radon transform [13]

Descriptively, every point of the Radon transform $RT_f(\phi, u)$ represents the summation of values of a function along a line $Z(t)$ at a distance u from the center and angle ϕ. The graphical representation of the Radon transform is also called Sinogram (as shown in Figure 3.3) as it appears like a number of sine waves with different phases and amplitudes. The x-axis of Sinogram represents angle ϕ at which a projection is computed. The y-axis represents distance u from the center. The important properties of the Radon transform are linearity, symmetry, periodicity, rotation by an angle, translation by a vector and scaling by a factor. These properties are defined below [14].

$$\text{Linearity} : RT_{(f+h)}(\phi, u) = RT_f(\phi, u) + RT_h(\phi, u) \tag{3.4}$$

$$\text{Periodicity} : RT_f(\phi, u) = RT_f(\phi \pm 2k\pi, u), k \in Z \tag{3.5}$$

$$\text{Symmetry} : RT_f(\phi, u) = RT_f(\phi \pm (2k+1)\pi, -u), k \in Z \tag{3.6}$$

$$\text{Translation by a vector} : h = T_(x_o, y_o)(f) \tag{3.7}$$

$$\implies RT_h(\phi, u) = RT_f(\phi, u - x_o\cos(\phi) - y_o\sin(\phi)) \tag{3.8}$$

$$\text{Rotation by an angle} : h = R_{\phi_o}(f) \implies RT_h(\phi, u) = RT_f(\phi - \phi_o, u) \tag{3.9}$$

$$\text{Scaling by a factor} : h = S_\alpha(f) \implies RT_h(\phi, u) = \alpha RT_f(\phi, u/\alpha), \alpha > 0 \tag{3.10}$$

By using these properties, the Radon transform can be used to compute the underlying transformation parameters relating two images or patterns directly from their projections only and without any additional computing [14]. To deal with the rotational misalignment of images in different image matching applications, the rotation property of RT can be used directly. The detailed explanation of the rotation property of the Radon transform is discussed below.

3.3.1 Rotation Property of Radon transform

The rotational property of the Radon transform states that rotation in the pixel domain is equivalent to a shift in the Radon domain. Let f_r represent the object image rotated by an angle ϕ_0 i.e.f_r=$R_{\phi_0}(f)$ where the rotation angle ϕ_0 can lie in the range $\phi_0 \in [1, 360]$. The rotational property of the 2π Radon transform states that the Radon transform at any point (ϕ, u) of a rotated image (rotated by angle ϕ_0) is equal to the Radon transform of the reference image at point $(\phi - \phi_0, u)$ [14].

$$f_r = R_{\phi_o}(f) \Longrightarrow RT_{f_r}(\phi, u) = RT_f(\phi - \phi_o, u) \tag{3.11}$$

Equation 3.11 simply defines the rotation property for all the rotation angle between $[1, 360]$. It shows that the rotation in a pixel domain is changed into a circular shift in the Radon domain. The size of the 2π Radon transform is $360 \times u$ where u represents the number of components in each Radon projection. The maximum value of u is equal to the length of the diagonal of an image. For an image of size $N \times N$, the value of u will be an integer value of $N\sqrt{2}$.

The Sinograms of two rotationally misaligned images are same except there is a circular shift along the direction component ϕ (as shown in Figure 3.3). This property is employed to estimate the rotation angle between two images using their Sinograms. Knowing the fact that rotation in the space domain is converted into a circular shift in the Radon domain, the Sinograms are compared to estimate the

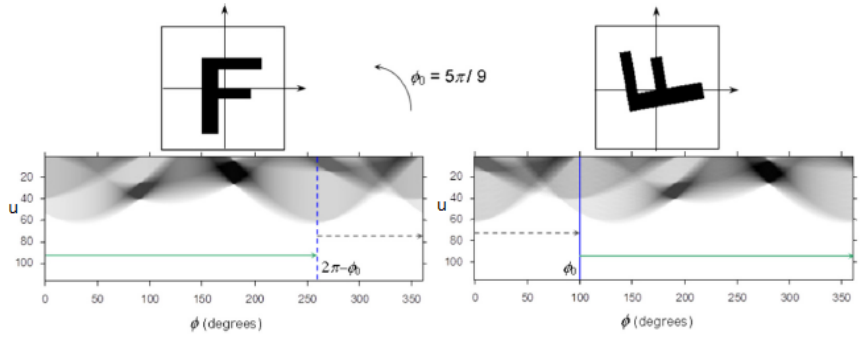


Figure 3.3: The 2π Radon transform rotation property: the rotation of an image by angle ϕ_0 is converted into the circular shift of ϕ_0 in the 2π Radon transform [14].

rotational angle between the images by using the equation given below:

$$F(\phi_0) = \int_0^{2\pi} \int_{-\infty}^{\infty} |RT_{f_r}(\phi, u) - RT_f(\phi - \phi_o, u)| du d\phi \tag{3.12}$$

The rotation angle ϕ_0 between $[1, 360]$ is given by the minimum value of F

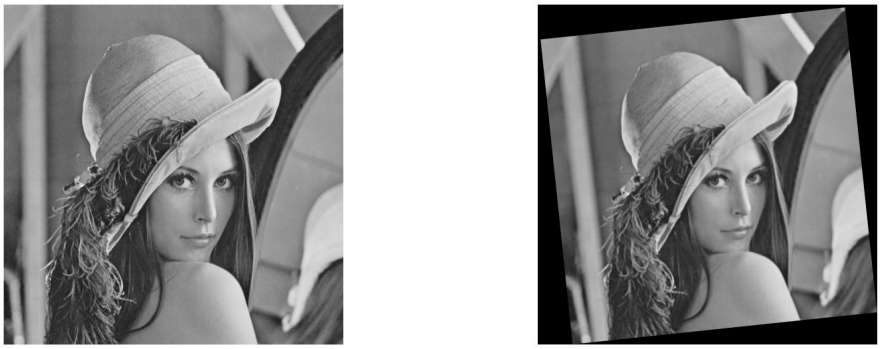
$$\phi_0^* = arg \min_{\phi_0} F \tag{3.13}$$

This property of the Radon transform can be exploited to solve the issue of a rotational difference between images. The details of using the Radon transform for computing the parameters are discussed in Section 3.4.

3.4 Method of Computing Parameters

As discussed earlier, alignment of the images is important for several image processing and computer vision applications. For alignment, we need to compute the parameters relating two images. However, the computation of parameters accurately and efficiently is still a challenging problem. In order to identify the parameters, we shall use the rotational property of the Radon transform for computing the angle, the ratio of the intensities for computing the scale, and the centroids of the images and the translation property of the Radon transform for computing

translation. Figure 3.4 shows two images of Lena: (a) Reference image I_R ; (b) Sensed image I_S obtained by using Equation 3.2.



(a) (b)

Figure 3.4: (a) Reference Lena Image (I_R); and (b) Sensed Lena image (I_S) obtained by applying the RST (Rotation, Scale and Translation) transformations [6, 0.9, 4, 4]

The idea is to compute transformation parameters one by one. The simplest parameter to compute is the scale between two images. The sum of the pixels in an image can be used to find the scale. The sums of pixels of reference and sensed images are given by Equation 3.14 and 3.15:

$$SumI_R = \sum_{j=1}^{N}\sum_{k=1}^{N} I_R(j,k) \tag{3.14}$$

$$SumI_S = \sum_{j=1}^{N}\sum_{k=1}^{N} I_S(j,k) \tag{3.15}$$

where $N \times N$ represents the size of the image. The scale α of the image is computed by:

$$\alpha = \sqrt{\frac{SumI_S}{SumI_R}} \tag{3.16}$$

Once the scale is computed, the next step is to find a common reference point between images. The most commonly used and easily computable reference point in an image is the centroid. The centroid of an image is the average location of all

the parts of an object inside an image. In other words, it is a position at which the weighted position vector of all parts of the object sum to zero [196]. For objects of uniform density like a circle, the centroid lies at the center of the object. However, for the objects with non-uniform shapes, it is a point where uniform force acts on object. The centroid is computed separately for each axis. Along x and y axis of the reference image, the centroid (C_{Rx}, C_{Ry}) is computed by using Equation 3.17 and 3.18 at any angle ϕ.

$$C_{Rx} = \frac{\int_{-\infty}^{\infty} uRT_R(\phi, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\cos(\phi) - \frac{\int_{-\infty}^{\infty} uRT_R(\phi + \pi/2, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\sin(\phi) \tag{3.17}$$

$$C_{Ry} = \frac{\int_{-\infty}^{\infty} uRT_R(\phi, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\sin(\phi) + \frac{\int_{-\infty}^{\infty} uRT_R(\phi + \pi/2, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\cos(\phi) \tag{3.18}$$

where, $RT_R(\phi, u)$ represents the Radon transform of the image I_R. Similarly, the centroid of the sensed image (C_{Sx}, C_{Sy}) can also be computed by using the Radon transform RT_S of the sensed image. Once the centroid of the object image is found, the next step is to shift the centre of the image to the centroid to have a common centre for both images. This eliminates the effect of the translation on the image, and consequently on the Radon transform. Alternatively, the Radon transform of the shifted reference and sensed images can be computed directly by using the translation property of the Radon transform and is given by Equations 3.19 and 3.20, respectively.

$$RT_{Rc}(\phi, u) = RT_R(\phi, u + C_{Rx}cos(\phi) + C_{Ry}sin(\phi)) \tag{3.19}$$

$$RT_{Sc}(\phi, u) = RT_S(\phi, u + C_{Sx}cos(\phi) + C_{Sy}sin(\phi)) \tag{3.20}$$

where RT_{Rc} and RT_{Sc} represent the Radon transforms of the reference and sensed images computed at the centroid. This process has eliminated the effect of trans-

lation on the Radon transform of the image. However, the Radon transforms R_{Oc} and R_{Tc} are still scaled and circularly shifted versions of each other.We know that the Radon transform shifts the rotation in the pixel domain to a shift in the Radon domain. Hence, this property can be used for computing the angle by comparing the Radon transforms for each possible angle [14]. However, comparing the Radon transform for each angle is computationally expensive.

As the sensed image is scaled, the Radon transform of the sensed image has α times the number of components to that of the reference image Radon transform. This is because the length of the diagonal of the sensed image is α times the length of the diagonal of the reference image; and the number of components in the Radon transform depends on the diagonal length. For computing the angle between both images, instead of comparing each component of the Radon transform, the consolidated information of the Radon transform for each angle can be compared. The common methods to consolidate the information for a set of values are mean and variance. The mean of a set of variables represents the central location of the distribution of values, whereas the variance measures the dispersion of the values around the mean. Moreover, the variance is a better perspective for a set of values as compared to the mean [197]. Hence, we will consolidate the information of each angle by a variance and represent the Radon transform matrix by the Row Variance Vector (RVV). The RVV_R of the reference image is computed as:

$$RVV_R(\phi) = var_u(RT_{Rc}(\phi, u)) \tag{3.21}$$

where RT_{Rc} represents the Radon transform of reference image computed at the centroid. Similarly, the Row Variance Vector of the sensed image represented by RVV_S can also be computed.

Scale Independent Row Variance Vector: The Row Variance Vector is scale independent as shown in Figure 3.5, because, it is computed by summarising

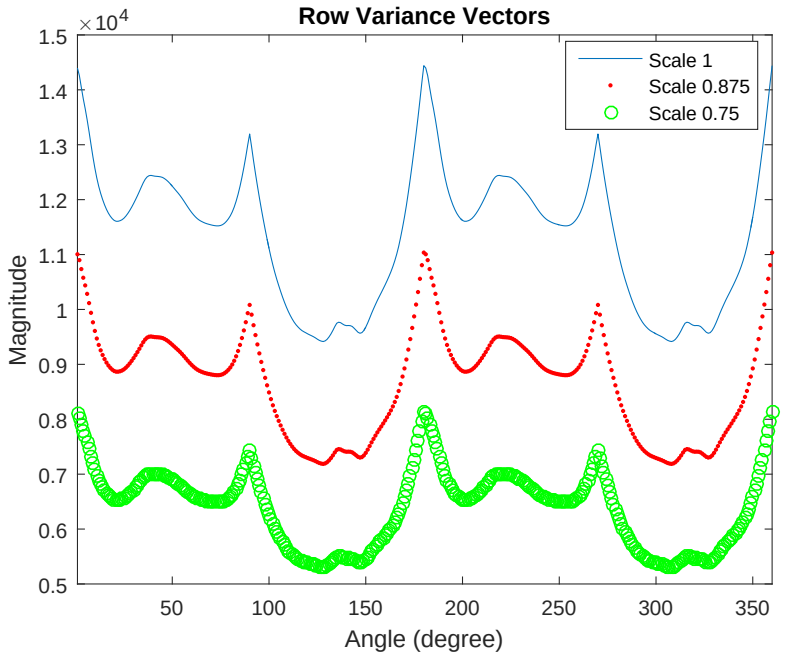


Figure 3.5: The Row Variance Vectors of the Lena image (shown in Figure 3.4 (a)) at different scales factors

the information of all components of the Radon transform at all angles individually. Therefore, even if the components are uniformly sampled due to scaling, it will result in similar Row Variance Vectors of different magnitudes as shown in Figure 3.5. After normalising the magnitudes of the vectors, the vectors of the scaled images are identical (as shown in Figure 3.6).

Figure 3.5 shows RVV of the Lena image at three different scales (1, 0.875, and 0.75). The Row Variance Vectors of the image at different scales are of the same length and are identical except the magnitude varies according to the scale. However, the normalised RVVs at different scales shown in Figure 3.6 are identical. Therefore, to compute the rotation angle, the normalised Row Variance Vectors of two scaled images can be compared directly unlike the method given in [14] where the Radon transforms of the reference and sensed images are scaled to have the same dimensions. Hence, in the proposed method, there is no need to have images with identical scaling.

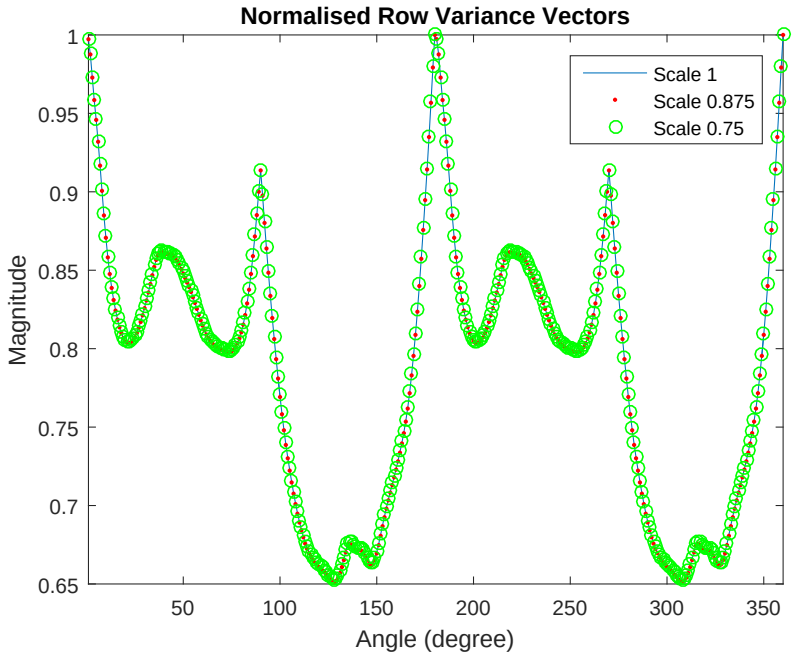


Figure 3.6: The normalised Row Variance Vectors of the Lena image (shown in Figure 3.4) at different scales

Moreover, the property of the Radon transform for representing the rotation in the pixel domain by a shift in the Radon domain is also preserved by the Row Variance Vector (as shown in Figure 3.7). In other words, the shifts between the Row Variance Vectors represent the rotational difference between images. Therefore, to increase the efficiency of the computation process, instead of using the complete Radon transforms to compute the angle of rotation, we can simply use the Row Variance Vectors (RVV) of both images without rescaling. Consider the following objective function:

$$H(\phi_0) = \int_0^{2\pi} |RVV_R(\phi) - RVV_S(\phi - \phi_0)| d\phi \tag{3.22}$$

Then, the rotation angle, ϕ_0^*, between angles 0 to 2π is given by:

$$\phi_0^* = arg(min_{\phi_0}(H)) \tag{3.23}$$

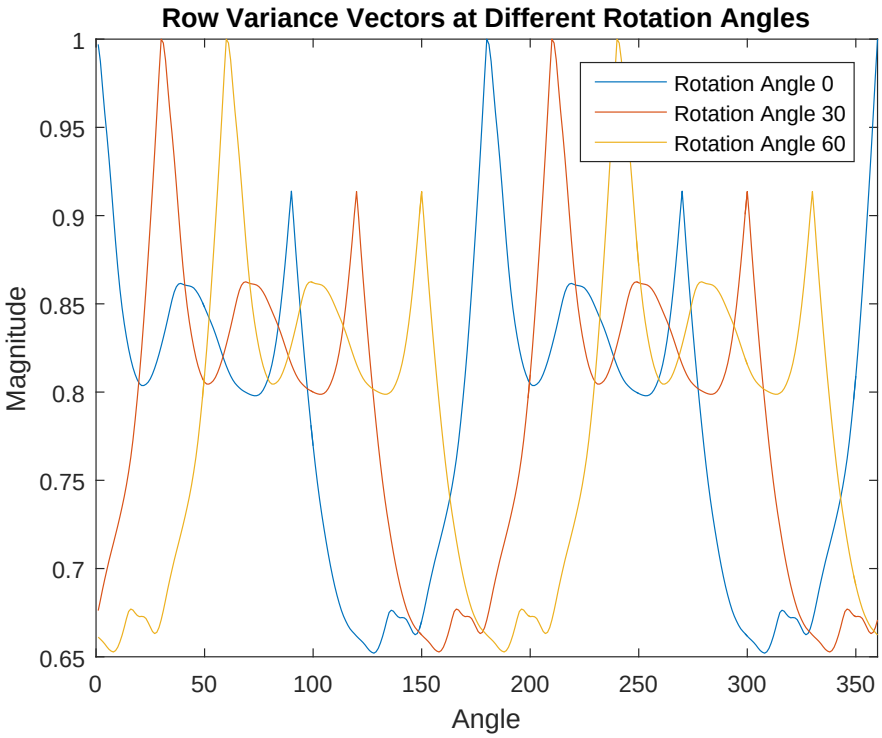


Figure 3.7: Row Variance Vectors of the Lena image (shown in Figure 3.4) at different rotation angles

The next parameter to estimate is the translation between two images. To find the translation, we used the scale (α), rotation angle (ϕ_0) and reference points (C_{Rx}, C_{Ry}) and (C_{Sx}, C_{Sy}) between two images by way of the centroid. The translation parameters (x_0, y_0) are computed by Equations 3.24 and 3.25, respectively.

$$y_0 = C_{Sy} + \alpha[C_{Rx}\sin(\phi_0^*)) - C_{Ry}\cos(\phi_0^*))] \tag{3.24}$$

$$x_0 = C_{Sx} - \alpha[C_{Rx}\cos(\phi_0^*)) + C_{Ry}\sin(\phi_0^*))] \tag{3.25}$$

The step by step procedure of computing the parameters is:

i) Compute the scale of the sensed image by using the ratio of means, using Equation 3.16.

ii) Compute the centroids of both images, using Equation 3.17 and 3.18.

iii) Compute the Radon transform around the centroid of the images, using Equation 3.19 and 3.20.

iv) Compute the Row Variance Vectors of the reference and sensed images, using Equation 3.21

v) Compute the objective function for computing the rotation angle, using Equation 3.22.

vi) Estimate the rotation angle, using Equation 3.23.

vii) Compute the translation coordinates, using Equation 3.24 and 3.25.

The algorithms of computing transformation parameters by using the method given in [14] and the proposed method is given in Algorithms 3.1 and 3.2, respectively.

3.5 Computational Time

The computational time of an approach plays an important role in matching images. One of the recent Radon transform based approaches proposed in [14] for computing the transformation parameters is highly accurate. However, the approach to computing the parameters is computationally complex. The algorithms for computing the parameters in [14] and the proposed method are Algorithm 3.1 and Algorithm 3.2, respectively. The computational times required for computing the scale and translation in both the methods are the same. However, the time differs when computing the rotation angle between the images.

In [14], to compute the rotational angle, the Radon transform of the image is requires translation, scaling and comparison for each shift between 1 to 360. The computational time of computing the rotation angle is based on the number of comparisons. The number of comparisons made to compare two Radon transforms

Algorithm 3.1 Compute RST parameters relating two images [14].

Input ← Reference Image I_R, Radon transform RT_R, Sensed Image I_S and Radon transform RT_S

Output ← Computation of RST Parameters (ϕ_0^*, α,(x_0, y_0))

Scale: Compute the scale α between images by using the ratio of Radon transforms.

Rotation: Computing the rotation angle need the centroid of the image and scale between images:

i. Compute the centroid of the reference image (C_{Rx}, C_{Ry}) and the sensed image (C_{Sx}, C_{Sy}).

ii. Compute the Radon transforms at the centroids of the images (RT_{Rc} and RT_{Sc}).

iii. Rescale the Radon transform of the reference image according to the computed scale (RT_{Rcs}).

iv. Compute angle:

$\phi_0 \leftarrow 1$

while $\phi_0 \neq 361$ **do**

  $RT_{Sc}^n \leftarrow \text{CircularShift}(RT_{Sc}, \phi_0)$

  for $\phi = 1 : 360$ **do**

    $L_R \leftarrow RT_{Rcs}(\phi, 1 : N\sqrt{2})$

    $L_S \leftarrow RT_{Sc}^n(\phi, 1 : N\sqrt{2})$

    $\text{Distance}(\phi) = \int_{u=1}^{N\sqrt{2}} (L_R(u) - L_S(u))du$

  end for

  $\text{F}(\phi_0) = \sum_{\phi=1}^{360} |Distance(\phi)|$

  $\phi_0 \leftarrow \phi_0 + 1$

end while

v. Take ϕ_0 as the rotation angle ϕ_0^* that results in the minimum value of F.

Translation: Using the computed scale, rotation angle and centroid, compute the translation parameters (x_0, y_0)**.**

are $P \times N\sqrt{2}$, where P represents the number of projections between 0 to 2π and $N \times N$ is the size of images. To find the angle, one Radon transform is shifted for each angle between 1 to 360 and compared with the other Radon transform. Hence, the total number of comparisons made are $P^2 \times N\sqrt{2}$.

The proposed approach converts a two dimensional Radon transform into a one dimensional vector of length P. The vectors are compared to compute the angle. For computing the rotation angle, the Row Variance Vectors are compared rather than the Radon transforms of images. The number of comparisons made to compare two vectors are P. To find the angle, one of the vectors is shifted for

Algorithm 3.2 Compute RST parameters relating two images using the proposed approach.

Input ← Reference Image I_R, Radon transform RT_R, Sensed Image I_S and Radon transform RT_S
Output ← Computation of RST Parameters (ϕ_0^*, α,(x_0, y_0))
Scale: Compute the scale α between images by using the ratio of pixel intensities of images.
Rotation: Computing the rotation angle needs only centroids of images:
i. Compute the centroid of the reference image (C_{Rx}, C_{Ry}) and the sensed image (C_{Sx}, C_{Sy}).
ii. Compute the Radon transforms at the centroid of the images (RT_{Rc} and RT_{Sc}).
iii. Compute the Row Variance Vectors of the images (RVV_R and RVV_S).
iv. Compute the angle:
for $\phi_0 = 1 : 360$ **do**
$\mathrm{H}(\phi_0) = \sum_{\phi=1}^{360} |RVV_R(\phi) - RVV_S(\phi - \phi_0)|$
end for
v. Take ϕ_0 as the rotation angle ϕ_0^* that results in the minimum value of H.
Translation: Using the computed scale, rotation angle and centroid, compute the translation parameters(x_0, y_0)**.**

each projection between 1 to 360 degrees and compared with other vector. Hence, the total number of comparisons made is P^2.

The computational time of the proposed approach shows that the proposed method is more efficient by a factor of $N\sqrt{2}$ compared to the approach in [14]. Not only is the proposed method more efficient, but the computation of the rotation angle is independent of the scale of the images. Whereas, in [14], to compute the rotation angle, each component of the matrices is compared independently. Therefore the Radon transforms need to be on the same scale. However, in the proposed approach the vector generated is of constant length for any scaled image. This property of the Row Variance Vector helps in improving the computational time of the proposed method.

3.6 Experimental Results

In order to demonstrate the effectiveness and accuracy of the computed parameters for the proposed method, experiments were conducted for computing the RST

parameters between two images which were corrupted or not corrupted by additive noise. The experiments were conducted to demonstrate schematically the various steps of the proposed approach for real world image applications. For this, we have chosen images from different applications such as the standard Lena and Barbara test images, a hand CT image [20] and a fingerprint image from FVC2002DB1 dataset. For the experiments, the dimensions of the Lena and Barbara images are 512×512 pixels, hand's Computed Tomography (CT) image is of size 285×285 and the fingerprint image is size 374×339. The fingerprint image is taken from FVC2002DB1 [3] database. These images can be divided into two types: the first, which are subjected to scale such as Lena or Barbara images; and the second, which are generally captured at same scale like fingerprints or CT images.

The first set of experiments is performed for the Lena and Barbara images with all three transformations namely, scale, rotation and translation. The second set of experiments are performed on the CT images and fingerprints under different rotation and translation. The algorithms are developed in a Matlab environment and the programs are run on the Intel Xeon 3.50 GHz processor and 64 GB of RAM.

Depending on the applications, the images are subjected to a number of RST (Rotation, Scale and Translation) transformations and corrupted by noise. The noisy images were obtained by adding noise to the reference image. The Gaussian white noise and salt and pepper noise are used for corrupting the images. The experiments are conducted by adding varying variance for Gaussian noise and density for salt and pepper noise. The noise is generated randomly, hence for statistical validity of the experiments, the experiments are performed 40 times and the average results are reported. The error for each of the computed transformation parameters is given, where the error is defined by the average difference between the given parameters and computed parameters.

$$Error = \frac{1}{K}\sum_{l=1}^{K} |P - P_l| \tag{3.26}$$

where, K represents number of times the experiment is performed, P represents real parameters $(\alpha, \phi_0, x_0, y_0)$ values and P_l represents the computed parameters (α_l, ϕ_{0l}, x_{0l}, y_{0l}). The error is computed for varying noise levels for both Gaussian and salt and pepper noises.

The performance of the first set of experiments on the Lena and Barbara images of size 512×512 is compared to the best performing algorithms proposed in [14] and [198]. The images are corrupted by adding Gaussian white noise with different SNR varying from ∞ to 2. The approach proposed in [14] is based on the Radon transform as well, however, it uses an exhaustive way of computing the angle. However, the approach discussed in [198] deals with the transformations in images by using the Radon transform, log mapping and the amplitude extraction. As shown in Table 3.1, the proposed approach performs equally well as compared to the approach in [198] in the case of scale and rotation, but it performs better for translation. The error reported for translation parameters in [198] is 1 (integer value) for both parameters, which means the error is rounded to an integer value. Therefore, the real value of the error can be even greater than the reported value. On the other hand, the accuracy of computing parameters for the proposed approach is as good as the approach in [14]. Whereas, the proposed approach can achieve the same accuracy in less computation time compared to the approach in [14].

To validate the proposed method, the accuracy of the estimated parameters is further checked by applying different transformations to the Lena and Barbara images. Tables 3.2 and 3.3 demonstrate the superior performance of the proposed method in computing the parameters under different transformations. The mean errors in scale, rotation and translation reported on the Lena and Barbara images

Methods	Noise (SNR)	Lena				Barbara			
		Scale	Rotation	Translation		Scale	Rotation	Translation	
		α	θ	X	Y	α	θ	X	Y
		0.9	6	4	4	0.9	6	4	4
Proposed Method	∞	0.0001	0	0.29	0.39	0.0001	0	0.40	0.36
	8	0.0004	0	0.42	0.29	0.0003	0	0.34	0.37
	2	0.0006	0	0.42	0.35	0.0005	0	0.39	0.50
Method in [14]	∞	0.0001	0	0.44	0.25	0.0001	0	0.41	0.35
	8	0.0004	0	0.43	0.23	0.0003	0	0.38	0.36
	2	0.0006	0	0.41	0.32	0.0005	0	0.42	0.45
Method in [198]	∞	0.0004	0	1	1	0.0004	0	1	1
	8	0.0004	0	1	1	0.0004	0	1	1
	2	0.0004	0	1	1	0.0004	0	1	1

Table 3.1: The errors in parameter estimation for the Lena and Barbara images by using the proposed approach at different SNR with Gaussian noise

Image	Real Parameters				Computed Parameters			
	Scale	Rotation	Translation		Scale	Rotation	Translation	
	α	θ	X	Y	α	θ	X	Y
Lena	0.9	6	4	4	0.9001	6	4.29	3.61
	0.9	6	10	7	0.9003	6	9.93	7.78
	0.9	12	9	9	0.9003	12	8.81	9.95
	0.8	192	9	14	0.8007	192	9.55	14.11
	1.3	353	15	15	1.3005	353	14.46	15.60
Mean Error					0.0004	0	0.31	0.56

Table 3.2: The estimated parameters for the Lena image by using the proposed approach as compared to the real parameters

Image	Real Parameters				Computed Parameters			
	Scale	Rotation	Translation		Scale	Rotation	Translation	
	α	θ	X	Y	α	θ	X	Y
Barbara	0.9	16	4	4	0.9003	16	4.55	3.62
	0.8	18	14	3	0.8006	18	13.44	3.67
	1.2	4	20	30	1.2010	4	20.15	29.81
	1	9	5	-5	1.000	9	4.78	-4.89
	0.9	8	8	5	0.9003	8	7.51	4.88
	0.9	9	15	-7	0.9003	9	15.32	-7.12
	0.8	10	-35	-27	0.8006	10	-34.93	-27.11
Mean Error					0.0004	0	0.3	0.24

Table 3.3: The estimated parameters for the Barbara image by using the proposed approach as compared to the real parameters

Noise	Barbara				Lena			
	Scale	Rotation	Translation		Scale	Rotation	Translation	
	α	θ	X	Y	α	θ	X	Y
	0.8	8	14	3	0.8	8	14	3
0.00	0.0003	0	0.58	0.50	0.0003	0	0.30	0.34
0.04	0.0003	0	0.55	0.20	0.0003	0	0.48	0.24
0.08	0.0004	0	0.51	0.20	0.0004	0	0.53	0.23
0.12	0.0005	0	0.57	0.18	0.0004	0	0.58	0.13
0.16	0.0006	0	0.47	0.38	0.0005	0	0.54	0.20
0.20	0.0006	0	0.53	0.31	0.0006	0	0.64	0.26

Table 3.4: The estimated parameters by using the proposed approach on the Barbara and Lena images at different densities of salt and pepper noise varying from 0 to 0.20 with an increment of 0.04 (Barbara images with varying densities is shown in Figure 3.8).

are (0.0004, 0, 0.31, 0.56) and (0.0004, 0, 0.30, 0.24), respectively. Moreover, the accuracy is also verified by adding salt and pepper noise to the images. The salt and pepper noise is added to the images with different noise densities varying from 0 to 0.20 with an increment of 0.04 (as shown in the Figure 3.4). The errors in computing parameters for salt and pepper noise are shown in Table 3.4. The accuracy of parameters computed is still maintained even at higher noise and for different transformations. The results for both Gaussian and salt and pepper noises demonstrate the robustness of the proposed method to different types of noises.

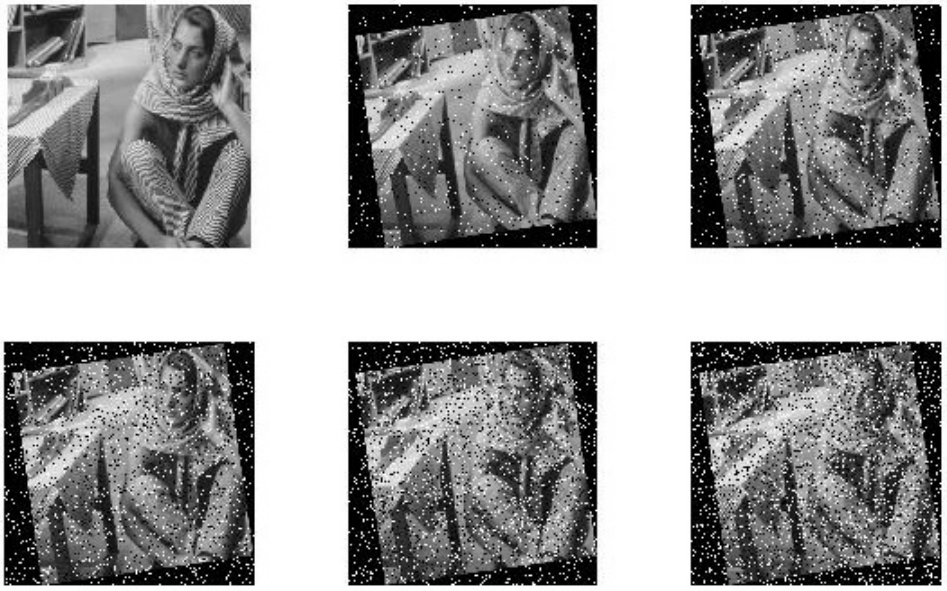


Figure 3.8: Barbara reference image (top left) and sensed images obtained by applying the RST transformations (8,0.8,14,3) and with varying salt and pepper noise density (d=0:0.04:0.20)

The second set of experiments is performed on the images which are generally captured at the same scale. For these experiments, hand CT images (shown in Figure 3.9) and fingerprint images (shown in Figure 3.10) are used. The performance is checked under both Gaussian and salt and pepper noise with different transformations. For Gaussian noise, the variance varies from 0 to 0.1 with an increment of 0.01, whereas, for salt and pepper noise, the noise density varies from 0 to 0.20 with an increment of 0.02. Table 3.5 and 3.6 shows the results of parameters computed for the hand CT images and the fingerprint images. For a hand CT image, the proposed method results in no error in the computation of angle under salt and pepper noise, whereas it results in a very small error of 0.025° at high Gaussian noise with a variance of more than 0.09. This has happened because the hand CT image is much smaller than the Lena and Barbara images and as noted in [199] and [11], the performance depends on the ratio of noise in image and size of image. Moreover, the error in the translation component is small as well. The mean errors in translation for the Hand CT image with salt and pepper

Image	Noise (Salt & Pepper)	Rotation	Translation		Noise (Gaussian)	Rotation	Translation	
		θ	X	Y		θ	X	Y
		16	19	14		38	9	4
Hand CT Image	0	0	0.26	0.26	0	0	0.32	0.54
	0.02	0	0.21	0.30	0.01	0	0.36	0.67
	0.04	0	0.19	0.26	0.02	0	0.41	0.62
	0.06	0	0.23	0.28	0.03	0	0.36	0.78
	0.08	0	0.22	0.38	0.04	0	0.43	0.76
	0.10	0	0.19	0.31	0.05	0	0.46	0.85
	0.12	0	0.22	0.39	0.06	0	0.53	0.88
	0.14	0	0.32	0.38	0.07	0	0.52	0.71
	0.16	0	0.29	0.41	0.08	0	0.62	0.97
	0.18	0	0.30	0.58	0.09	0.025	0.64	0.94
	0.20	0	0.45	0.78	0.1	0.025	0.58	0.85

Table 3.5: The errors in parameter estimation for the hand computed tomography image (shown in Figure 3.9) at different levels of both salt and pepper and Gaussian noise

noise and Gaussian noise are (0.26,0.39) and (0.47,0.79), respectively. In the case of the fingerprint, the proposed method results in no error in the computation of the angle under both noises and reports small errors in translation components and estimated parameters are close to the real parameters. The mean errors in translation of the fingerprint with salt and pepper noise and Gaussian noise are (0.27,0.45) and (0.28,0.58), respectively.

The results discussed above for different types of images demonstrate the accuracy and robustness of the proposed method of computing parameters. The proposed method of computing parameters results in very small errors for images from different applications. Moreover, the computed parameters are very close to real values of parameters applied to the reference image with or without noise. Not only is the proposed method accurate and robust to noise, but also it computes the parameters faster than the best performing algorithms.

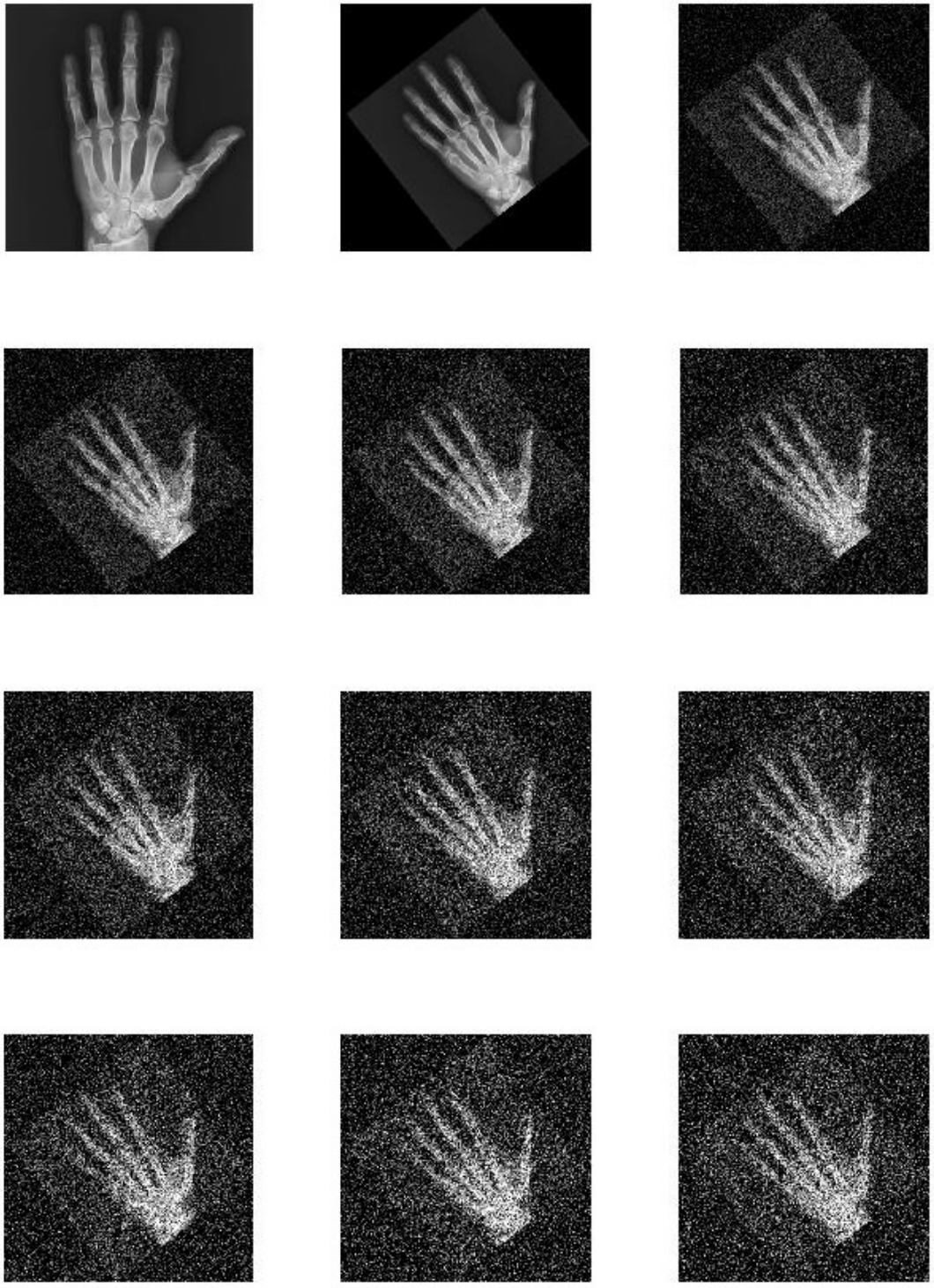


Figure 3.9: Hand CT reference image (top left) and sensed images obtained by applying Rotation and Translation transformations (38,9,4) and Gaussian noise with varying variance from 0:0.01:0.1

Image	Noise (Salt & Pepper)	Rotation	Translation		Noise (Gaussian)	Rotation	Translation	
		theta	X	Y		θ	X	Y
		17	5	20		17	15	-16
Fingerprint	0	0	0.29	0.41	0	0	0.37	0.44
	0.02	0	0.29	0.40	0.01	0	0.37	0.45
	0.04	0	0.29	0.38	0.02	0	0.32	0.51
	0.06	0	0.26	0.46	0.03	0	0.34	0.49
	0.08	0	0.27	0.42	0.04	0	0.24	0.76
	0.10	0	0.28	0.41	0.05	0	0.24	0.74
	0.12	0	0.27	0.46	0.06	0	0.23	0.73
	0.14	0	0.27	0.39	0.07	0	0.32	0.44
	0.16	0	0.29	0.43	0.08	0	0.21	0.73
	0.18	0	0.26	0.58	0.09	0	0.23	0.68
	0.20	0	0.24	0.62	0.1	0	0.27	0.42

Table 3.6: The errors in parameters estimation for the fingerprint image (shown in Figure 3.10) at different levels of both salt and pepper and Gaussian noise

Figure 3.10: Fingerprint reference image and sensed image obtained by applying Rotation and Translation transformations (17,15,-16) and Gaussian Noise with variance 0.1

3.7 Conclusion

In this chapter, we have discussed a pixel based parameter computation method for images having pixel-to-pixel correspondence by using the Radon projections. To compute the rotation angle efficiently as compared to existing techniques, a Row Variance Vector based on Radon transform is proposed. The Row Variance Vector is scale invariant and this property helps in improving the efficiency of the method as images are not required to rescale for computing other parameters. Moreover, in low quality or noisy images, the rescaling of images can add errors in other parameter computation which we have eliminated by using the Row Variance Vector.

As the method is pixel based, the method is applicable to images of different applications. Therefore, the effectiveness of the proposed method is tested on images of varying classes and different resolution and results are compared with the existing methods. The parameters computed by the proposed method are very close to the parameters applied to the reference image. The errors reported by the proposed method in computing parameters are trivial and lower than those obtained by recent works on the same images. Moreover, the running time for the computation of the angle (that affects the computation time of the whole process) is low enough as compared to existing techniques. The proposed method can compute the parameters with better accuracy and in less computation time. The robustness of the proposed method to noise is also tested by adding Gaussian and salt and pepper noises. The results shows that high accuracy is maintained even when images are contaminated by noise.

The next chapter presents an efficient technique for matching images having pixel-to-pixel correspondence without aligning them. It also discusses the method of further improving the efficiency by reducing the number of projections and compares the results with other known techniques.

Chapter 4

Alignment-free Approach for Image Matching Applications

4.1 Preamble

Chapters 2 and 3 discussed that alignment of images is one of the crucial steps in most computer vision and image processing applications [24]. Alignment of images is a process of transforming a sensed image into the coordinate system of the reference image. Generally, the methods used for image alignment can be classified into different categories: pixel based and feature based [200]. Pixel based techniques either use intensity patterns of pixels by means of correlation or find the transformation parameters between sensed and reference images using different transforms such as Fourier and Radon, whereas feature based approaches identify the correspondence amongst extracted features of images.

Using the correlation as a measure in pixel based approaches is one of the most common methods for image alignment [17]. The correlation between two images is greater if and only if pixel-to-pixel correspondence is large. However, when correlation is used for aligning images, in order to find the best pixel-to-pixel correspondence between sensed and reference images, the sensed image needs to

be rotated by each possible angle and compared with the reference image [10]. This is an exhaustive approach that makes it inefficient for real time applications. Similarly, the transform based method proposed in [14] use Radon projections to match the images by aligning them using computed parameters. However, the complete process of computing parameters is computationally exhaustive.

Another approach to aligning images comprises the feature-based approaches which consist of mainly three steps: (i) extraction of reliable features in images; (ii) feature matching; and (iii) alignment of the images based on matched features [25]. Many methods for extracting the features have been developed in the literature [9]. For instance, in object recognition, feature like corners or lines are detected by using different approaches discussed in [2] [31] [27] [32], whereas in fingerprint recognition singularities or minutiae points are detected [3]. Matched features are used as reference points in both images for aligning them. However, there are issues related to these approaches such as (i) accurate detection of reliable features is very difficult; (ii) detection of features is time consuming; (iii) partial images may not consist of enough features; and (iv) extensive image processing is required. These issues complicate the process of aligning images and affect the efficiency of image matching processes.

As discussed above, researchers have been working to resolve the issue of image alignment in many fields such as remote sensing, recognition, panorama and registration. However, aligning images accurately and efficiently is still a challenging issue that affects the overall efficiency and accuracy of the matching process. Rotational and translational differences between images are the major causes of misalignment of images. The translational difference between images can be dealt with by using the centroids of images [14], but managing the rotational difference efficiently and accurately is difficult. Chapter 3 proposes a method of computing parameters efficiently that can be used for efficient image matching. However, despite reduced computation time for the estimation of parameters the overall

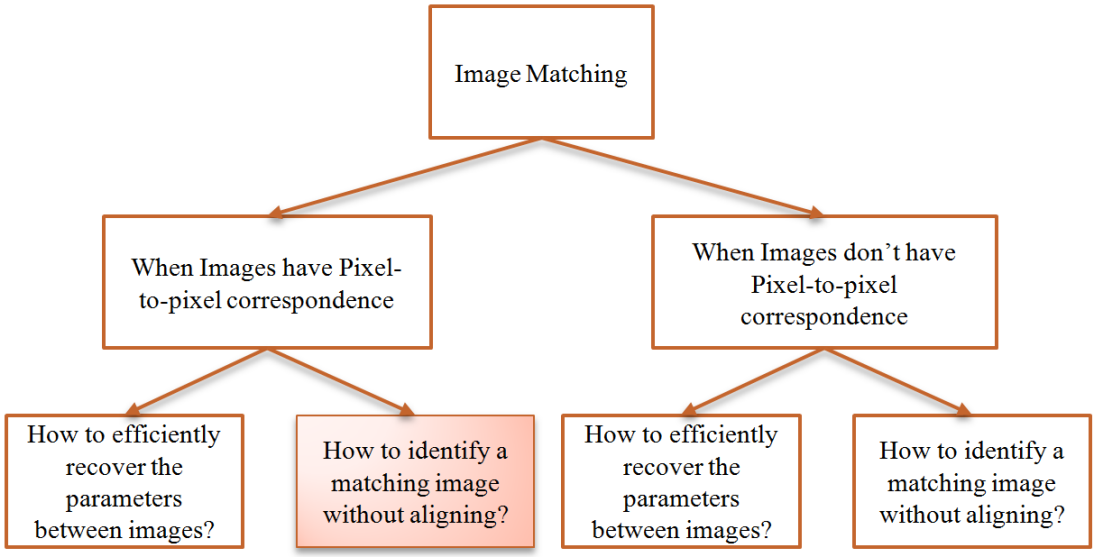


Figure 4.1: The coloured block represents the thesis objective addressed in this chapter.

total computation time for the matching is enhanced. Moreover, in recognition applications such as object recognition and biometric recognition, the alignment process is only an intermediate step. Therefore, it is not mandatory to align the images before matching. Hence, the question arises: is it possible to match images without computing parameters or aligning them? In other words, a method is required that can take care of the transformational differences between images without computing parameters and match images accurately and efficiently. This chapter proposes an alignment-free image matching method for images having exact pixel-to-pixel correspondence (shown in Figure 4.1) that not only eliminates the need for rotationally aligning the images but also improves the computational time of the matching process.

The remainder of the chapter is structured as follows: Section 4.2 discusses the need for a rotation invariant representation of an image for alignment-free matching. Section 4.3 discusses exploiting the rotation invariant property of the Radon transform, to compute the rotation invariant representation and deal with the rotational difference of images. Section 4.4 is dedicated to the proposed alignment-free

matching method. The computational time of the proposed approach is discussed in detail in Section 4.5. The performance of the proposed approach is demonstrated by applying it to the UMD Logo dataset in the Section 4.6. Finally, the conclusion is given in Section 4.7.

4.2 Need for a Rotation Invariant Representation

As we have discussed in the previous section, the major causes of mis-alignment between images are rotation and translation differences. The translation difference can be dealt with simply by finding the centroid of the images [10]. However, finding the rotational difference between the images is a very difficult and time consuming process. Therefore, if the images can be represented in the rotation invariant representation before matching, then there is no need to find the rotational difference between images.

To represent the images in a rotation invariant representation, we transform the image into the Radon domain. The images can be transformed into the Radon domain by applying the Radon transform. The Radon domain simplifies the rotational difference between images. The Radon transform converts the rotation in the pixel domain to a shift in the Radon domain. Figure 4.3 shows that the Sinograms of two rotationally unaligned but identical images (shown in Figure 4.2) are identical except there exists a shift along the direction ϕ. In other words, for images that differ only by a rotation, the Sinograms of those images are shifted versions of each other. Hence, to match these images, the Sinogram of sensed image can be compared with all the shifted Sinogram of the reference image. Comparison of the Sinograms helps in finding the maximum similarity between images. One of possible ways of comparing two Sinograms is by computing the Euclidean distance

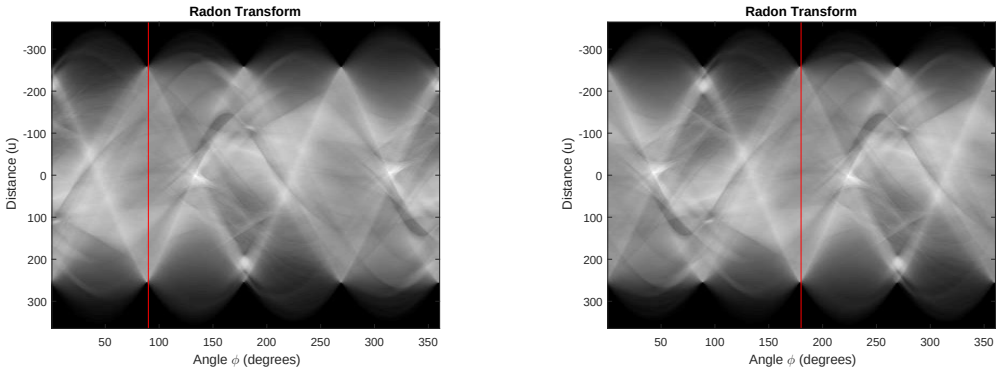
(a) (b)

Figure 4.2: Lena reference and 90° rotated images

between them [163]. The smaller the distance between Sinograms, the greater the similarity between images.

The use of Sinograms of images replaces the need to rotate the sensed image for each angle as is done in [10] by the need to shift the Sinogram along the direction component ϕ for each angle. Although, the Radon transform has simplified the process of matching images, it is not rotation invariant as the Sinogram needs to be shifted for each possible angle and comparison needs to be made. Therefore, one needs to get rid of the measurement of shift in the Sinograms and convert the pixel information of an image into a rotation invariant representation.

The shift in the Sinograms occurs along the direction component ϕ (as shown in Figure 4.3). To represent the image's pixel information in a rotation invariant form, shift needs to be eliminated. Then, instead of comparing information at each point in the Sinograms, the consolidated information for each distance can be compared. By consolidating information along the direction component we can eliminate the need to shift the Sinograms. Consequently, shift independence in the Radon domain represents rotation independence in the pixel domain. This process of consolidating information and hence representing an image in a rotation invariant representation is described in the next section.



(a) (b)

Figure 4.3: The Sinogram of the Lena reference and 90° rotated images (the rotated image Sinogram is the shifted version of the reference image Sinogram).

4.3 Rotation Invariant Representation

The rotation independent representation of an image can be represented by a vector based on the Radon transform of an image [201]. The image is converted into a rotation invariant Variance Vector that is computed by using the Radon transform. The Radon transform is a matrix of projections of an image at different angles ranging from 1 to 360 degrees. A projection is computed by summing the pixel values along each column of the image at a certain angle ϕ. Hence, in a 2π Radon transform, there are 360 projections. The maximum number of components in each projection will be equal to the length of the diagonal of the image. Therefore, it results in a matrix of size $360 \times N\sqrt{2}$, where N represents the size of rows and columns of an image. Each projection represents a row in the Radon matrix as shown below:

$$RT_R(\phi, u) = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1N\sqrt{2}} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2N\sqrt{2}} \\ A_{31} & A_{32} & A_{33} & \dots & A_{3N\sqrt{2}} \\ \hline A_{(P-1)1} & A_{(P-1)2} & A_{(P-1)3} & \dots & A_{(P-1)N\sqrt{2}} \\ A_{P1} & A_{P2} & A_{P3} & \dots & A_{PN\sqrt{2}} \end{bmatrix} \tag{4.1}$$

where P represents the number of projections from 1 to 360 degrees. Matrices respresented by Equations 4.1 and 4.2 show an example of the Radon transform matrices RT_R and RT_{rot} of reference image I_R and ϕ_0 degrees rotated reference image I_{rot}, respectively. Comparing the two matrices column-wise will reveal the same data, but in the RT_{rot} matrix's column, the data values are circularly shifted. The shift is equal to the rotational difference between the images [195]. This property is true if and only if images exhibit pixel-to-pixel correspondence.

$$RT_{rot}(\phi, u) = \begin{bmatrix} A_{(P-(\phi_0+1))1} & A_{(P-(\phi_0+1))2} & \dots & A_{(P-(\phi_0+1))N\sqrt{2}} \\ \hline A_{P1} & A_{P2} & \dots & A_{PN\sqrt{2}} \\ A_{11} & A_{12} & \dots & A_{1N\sqrt{2}} \\ A_{21} & A_{22} & \dots & A_{2N\sqrt{2}} \\ \hline A_{(P-\phi_0)1} & A_{(P-\phi_0)2} & \dots & A_{(P-\phi_0)N\sqrt{2}} \end{bmatrix} \tag{4.2}$$

Since, we are interested only in determining whether two images are similar or not, there is no need to compute the defined shift. All we need to know is whether the set of values of the column pairs in matrices are identical. Given the nature of images, it is unlikely that the data values in the respective columns in the

two matrices will be identical. Instead, we would like to compute how close the corresponding column pairs of matrices are.

To eliminate the shift in the Radon domain, the column data in the matrix needs to be replaced by a single value that captures the properties of the data and is independent of the positions of individual values. Two simple metrics that can be used to represent a set of values are the arithmetic mean and variance. The arithmetic mean measures the central location of the distribution of values whereas the variance measures the dispersion of the values around the mean. Variance is a better perspective for a set of values compared to the mean [197]. One reason is that even when the values in a data set have huge differences, the variance can still contain meaningful information about the data [197]. Moreover, variances of different datasets tend to differ even when the means are similar. Therefore, we propose to replace the values along each column by the variance of the column values. This results in a vector of variances for the Radon transform matrix and this vector is referred as the Variance Vector in further discussion.

The columns of matrices shown by Equations 4.1 and 4.2 consist of identical data, except the data values are circularly shifted. Therefore the variances along corresponding columns will be identical. Equations 4.3 and 4.4 represent the Variance Vector of the reference and rotated images.

$$V_R(u) = var_\phi(RT_R(\phi, u)) \tag{4.3}$$

$$V_rot(u) = var_\phi(RT_{rot}(\phi, u)) \tag{4.4}$$

where, $V_R(u)$ and $V_{rot}(u)$ represent the Variance Vectors of the reference and rotated images and $RT_{rot}(\phi, u) = RT_R(\phi - \phi_0, u)$, where ϕ_0 is the angle of rotation. Since the data represented by $RT_R(\phi, u)$ and $RT_{rot}(\phi, u)$ at any distance u (along the corresponding column) is identical, therefore the Variance Vectors of the reference and rotated images are identical. The final representation of the image

is dependent only on one variable, namely the distance u, and is independent of angle ϕ. Hence it is rotation invariant.

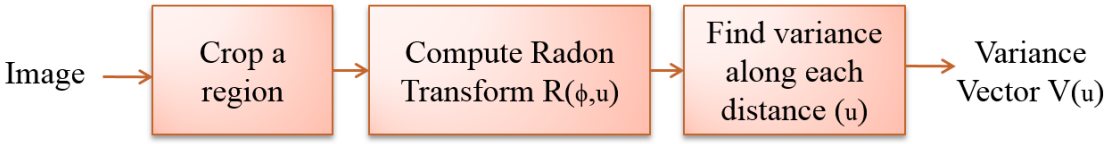


Figure 4.4: The block diagram of computing the rotation invariant Variance Vector

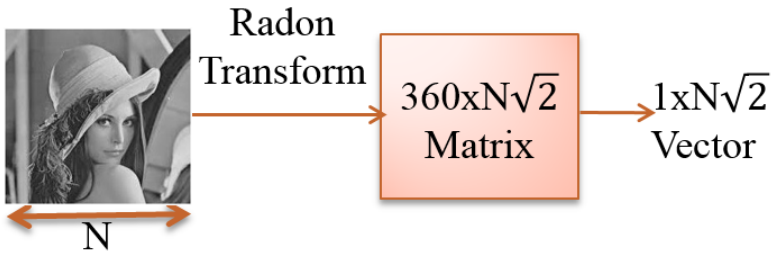


Figure 4.5: The process of converting an image in the pixel domain to a rotation invariant Variance Vector

The block diagram of computing the Variance Vector of an image is shown in Figure 4.4. The Variance Vector is not only rotation invariant but also preserves the fundamental attributes of the image's pixel information. Figure 4.5 shows the Lena image that is used as an example for computing the Variance Vector. For this example, the Radon transform projections are computed for each angle between 1 to 360 degrees resulting in a matrix of size $360 \times N\sqrt{2}$, where $N \times N$ is the size of the image. Computing the variance for each column results in a $N\sqrt{2}$ dimensional rotation invariant Variance Vector.

To demonstrate the rotation invariance of the computed vectors we have computed the Variance Vectors for the Lena image and its rotated equivalent (as shown in Figure 4.2). The normalised vectors shown in Figure 4.6 demonstrate that two Variance Vectors coincide perfectly. Besides rotation invariance, the Variance Vector is symmetric as well (as shown in Figure 4.6). Equation 4.5 shows that the

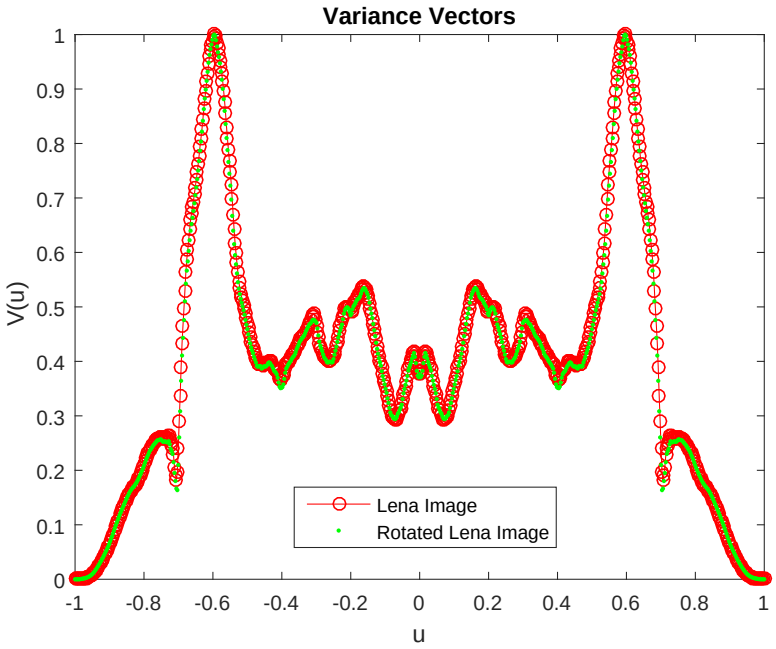


Figure 4.6: The normalised Variance Vectors of the reference and rotated Lena images (shown in Figure 4.2)

values of components on either side of the center are identical.

$$V_R(-u) = V_R(u) \tag{4.5}$$

Hence, a two dimensional image can be represented by a one dimensional symmetric rotation invariant Variance Vector. Only half of the components of the vectors need to be compared for matching images. As a consequence, the symmetric rotation invariant Variance Vector can be used for matching the images without rotationally aligning them.

4.4 Image Matching Algorithm

As discussed earlier alignment is a crucial step in different image matching applications. Among the existing methods of alignment, one of the popular approaches for aligning the images uses the computed underlying parameters relating two

images. However, computing the parameters for aligning images increases the computational complexity of the entire matching process. Therefore, if it is possible to match the images without computing all the parameters, then we will have an efficient matching algorithm.

Generally, the images are related to each other by different transformations namely, scale, translation and rotation. The existing methods used for matching the images have mainly three steps: extract the parameters, align the images using these parameters, and match the aligned images. The alignment of images makes the entire matching process computationally exhaustive. Therefore, we have developed a method that matches the images without aligning them. In other words, images are matched efficiently without retrieving all the transformation parameters.

The algorithm for matching the images depends only on the scale difference between images and the centroids of the images. The proposed approach does not require extraction of the rotational difference between images. Let us suppose we have two images I_R and I_S, where I_S is obtained by transforming the reference image I_R with scale α, rotation ϕ_0 and translation (x_0, y_0). The scale of the images can be computed by using the sum of the pixel values of the images. Equations 4.6 and Equation 4.7 represent the sum of the pixel values of the images. The scale relating the two images can be computed by using Equation 4.8.

$$SumI_R = \sum_{j=1}^{N} \sum_{k=1}^{N} I_R(j,k) \tag{4.6}$$

$$SumI_S = \sum_{j=1}^{N} \sum_{k=1}^{N} I_S(j,k) \tag{4.7}$$

$$\alpha = \sqrt{\frac{SumI_S}{SumI_R}} \tag{4.8}$$

Moreover, if only the Radon projections are given then the scale can be computed by using the Radon projections as is computed in [163]. The next parameter to compute for matching the images is the centroid of the images, which can be computed directly by using the Radon projections [163]. Let us suppose the Radon transforms of the images I_R and I_S are $RT_R(\phi, u)$ and $RT_S(\phi, u)$, respectively. The x and y coordinates of the centroid of reference image are represented by C_{Rx} and C_{Ry}, respectively. For any angle ϕ, the coordinates of the centroid (C_{Rx}, C_{Ry}) are computed as:

$$C_{Rx} = \frac{\int_{-\infty}^{\infty} uRT_R(\phi, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\cos(\phi) - \frac{\int_{-\infty}^{\infty} uRT_R(\phi + \pi/2, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\sin(\phi) \tag{4.9}$$

$$C_{Ry} = \frac{\int_{-\infty}^{\infty} uRT_R(\phi, u)dp}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\sin(\phi) + \frac{\int_{-\infty}^{\infty} uRT_R(\phi + \pi/2, u)du}{\int_{-\infty}^{\infty} RT_R(\phi, u)du}\cos(\phi) \tag{4.10}$$

The coordinates of the centroid is computed for different values of ϕ and final result is the average of all values. Similarly, the centroid of the sensed image (C_{Sx}, C_{Sy}) can be computed as well. Once the centroids of images are computed then the centres of the images are shifted to the centroids and the Radon transforms of the shifted images are computed. Then the origins of the system axes of both images will be at the same location. Otherwise, the Radon transform of the shifted images can be computed directly by using the translation property of the Radon transform [195]. The Radon transform of the shifted image at any angle ϕ is computed as:

$$RT_{Rc}(\phi, u) = RT_R(\phi, u + C_{Rx}\cos(\phi) + C_{Ry}\sin(\phi)) \tag{4.11}$$

$$RT_{Sc}(\phi, u) = RT_S(\phi, u + C_{Sx}\cos(\phi) + C_{Sy}\sin(\phi)) \tag{4.12}$$

where RT_{Rc} and RT_{Sc} represent the Radon transforms of center shifted versions of images I_R and I_S. As the Radon transforms are computed with the centroids, the effect of translation is eliminated. However, the Radon transform RT_{Sc} is still

a scaled and circularly shifted version of RT_{Rc} because the sensed image is scaled and rotated. Therefore, for matching the images, the Radon transform of the reference image needs to be scaled to the Radon transform of the sensed image by using the scaling property of the Radon transform. The reference image's Radon transform RT_{Rc} is scaled by using Equation 4.13 to have the same size as that of RT_{Sc}.

$$RT_{Rcs}(\phi, u) = \alpha RT_{Rc}(\phi, u/\alpha) \tag{4.13}$$

where RT_{Rcs} represents the scaled Radon transform of image I_R. By using the scaling and translation property of the Radon transform, the effects of scale and translation are eliminated and the Radon transforms RT_{Rcs} and RT_{Sc} are similar except being circularly shifted versions of each other.

Knowing the rotation in the space domain is converted into the shift in the transformed domain, the Radon transform can be compared to compute the angle for rotationally aligning them as is done in [14]. Once the images are aligned, the Euclidean Distance between the Radon transforms can be computed to match the images. However, this process needs to shift one of the Radon transforms for each angle from 1 to 360 degrees to compare it with the other Radon transform. This process of computation of angle affects the efficiency of the matching process. This raises the question of whether it is possible to match the images without computing the rotation angle?

We have discussed that two rotationally shifted Radon transforms can be compared by computing the rotation invariant Variance Vector instead of directly comparing the Radon transforms. These vectors can be used to compare the images without computing the rotation parameter. The Radon transforms R_{Rcs} and R_{Sc} are used to compute the Variance Vectors V_R and V_S of images I_R and I_S by using Equations 4.14 and 4.15, respectively.

$$V_R(u) = var_\phi(RT_{Rcs}(\phi, u)) \tag{4.14}$$

$$V_S(u) = var_\phi(RT_{Sc}(\phi, u)) \tag{4.15}$$

For computing the similarity between images, different similarity metrics can be applied to Variance Vectors. One of the commonly used metrics is the Euclidean distance that is computed by using Equation 4.16. The images are compared by computing the distance between the Variance Vectors. Ideally, for two identical but rotationally unaligned images, the two vectors will be same. Hence, the distance between these vectors will be zero. However, due to noise in images, the distance between the two vectors results in a non-zero value. Hence, the image pairs resulting in the smallest distance value can be considered as similar images, in matching applications.

$$\text{Euclidean Distance} = \sqrt{\sum_{u=1}^{N\sqrt{2}} (V_R(u) - V_S(u))^2} \tag{4.16}$$

The proposed method requires only the scale and centroid of images for matching them. Once, the scale and centroid of images are known, images are matched without computing translation and rotation parameters. By using the proposed method, it is possible to match images without aligning them. The algorithm for alignment-free image matching:

i) Compute the scale between the images by using Equation 4.8.

ii) Compute the centroids of both images by using the Equations 4.9 and 4.10.

iii) Compute the Radon transforms of both images around the centroids of images by using Equations 4.11 and 4.12.

iv) Rescale the Radon transform of reference image by using the Equation 4.13.

v) Compute the Variance Vectors of both images using Equation 4.14 and 4.15.

vi) Compute the Euclidean Distance between images using Equation 4.16.

vii) Select an image as matching image that results in minimum Euclidean distance with sensed image.

4.5 Computational Time

A Radon transform based object recognition method namely R-RST is proposed in [14]. The approach is accurate but computationally expensive as the major part of computation time is spent in minimizing the similarity measure to compute the rotation angle between two images. However, the proposed method has reduced the computational time of matching process by matching images without the need of computing the rotation angle between them.

Both the proposed and R-RST methods spend equal time in computing the Radon transform, scale and centroid of the images which is mandatory for both methods. However, computation of the rotation parameter between two images makes the entire difference. In R-RST, the rotation angle is used for computing the translation between images and aligning the images. The rotation angle between the images is computed by comparing the Radon transforms for each possible angle from 1 to 360 degrees. That is an exhaustive process. Also, for higher accuracy of a rotation angle, a larger number of projections need to be taken. That in turn, further increases the computation time of matching the images. Hence, in R-RST, the computational time of the matching process depends on the time required for computing the rotational angle. The number of computation steps for computing the angle is $(P^2 \times N\sqrt{2})$, where $N \times N$ is the size of the image and P represents the number of projections from angle 1 to 360 degrees. Hence, the computational time increases with the number of projections.

On the other hand, the computational time of the proposed matching process depends only on computation and comparison of Variance Vectors. The proposed method does not require computation of the rotational angle as the Variance Vector is rotation invariant. Unlike Radon transforms compared in[14], there is no need to compare the vectors by shifting for each possible angle. Therefore, the number of computation steps in the proposed method is only $N\sqrt{2}$ (equal to the length Variance Vector) and is independent of the number of projections. The proposed method takes P^2 times fewer steps for matching images as compared to R-RST. Hence, the proposed method matches the images faster than R-RST. Moreover, whenever the size of the image (N) is less than the number of projections (P), the proposed method matches the images faster than R-RST.

4.6 Performance Evaluation

As shown in Section 4.5, the computation time of proposed method required for matching the images is less than the computation time required for the R-RST method. However, the R-RST method performs best with high precision even at different high noise densities among all the region based descriptors, namely: Hu moments [202], Radon composite features (RCF)[203], Generic Fourier Descriptors (GFD) [204], Zernike moments [205] and wavelet ϕ R-signature (WϕR)[163] in terms of Precision-Recall [14]. Hence, it is important to verify the effectiveness of the proposed methods in terms of Precision-Recall as compared to the R-RST approach.

First, we assume that the image dataset used for experiments consists of isolated objects as is used in the R-RST method [206]. The logo dataset used in [14] is used for testing the effectiveness of the proposed approach, as the logos are popular examples of images having isolated objects. Moreover, the logo recognition has great importance in many domains and one of the important domains where logo

recognition is crucial is document identification [15]. By recognising the logos, the semantic information of the document is obtained. That helps in making a high level decision about the importance of the information in the textual components and whether it needs to be analyzed further. Logo recognition can be used in two different applications: first, given a logo and a logo database, we can determine if the logo exists in the database; second, given a logo and a document database, we can determine if there is any document in the database containing the given logo, which helps in finding the relevant documents efficiently. Moreover, logo recognition plays an important role in other applications such as vehicle logo detection for traffic control systems [207], augmented reality [208] and copyright infringement detection.

Logos are generally a combination of graphic icons and mixed text, which when identified activates the association of the graphic icon in the logo to an organisation or group. Using the proposed approach of image matching, experiments are conducted on the UMD logo dataset [15]. The dataset consists of 80 images of 10 different categories, each consisting of 8 images. These 8 images are generated by random scaling, translating, and rotating the corresponding logo images. Moreover, the robustness to noise of the proposed approach is tested by adding noise of different density to the images. The salt and pepper noise is added to corrupt the images. The noise is added with a different density 'd' representing the percentage of pixels corrupted in the image. The density of the noise ranges from 0.00 to 0.20 with an increment of 0.04. The left side of Figure 4.7 shows 10 logo images (one from each category). The right side of Figure 4.7 demonstrates the images with different transformation and noise densities. Moreover, the R-RST method also uses the same dataset, enabling comparison.

The criteria used for comparing the proposed method with the existing approaches is Precision-Recall. It is one of the commonly used metrics in object recognition, document identification, pattern recognition, information retrieval and

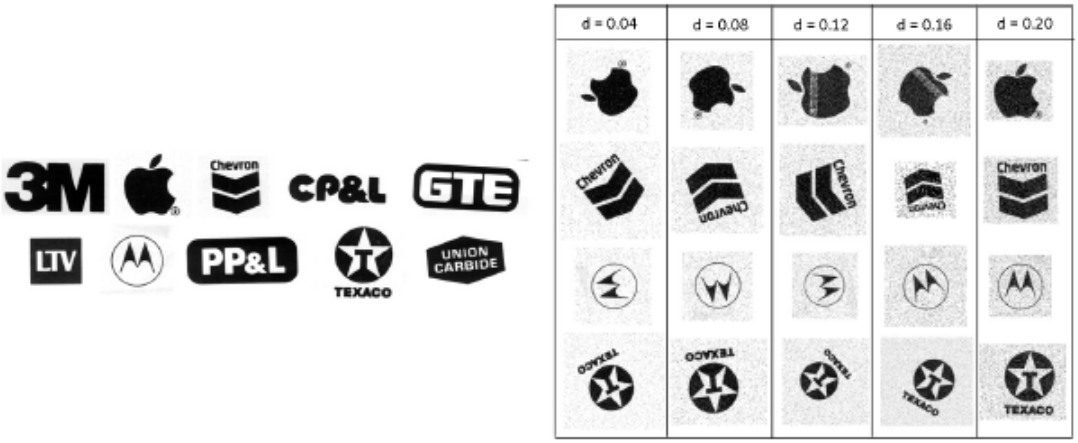


Figure 4.7: Ten noise-free logo images from the UMD dataset [15] and the samples of noisy images generated by adding the salt and pepper noise with densities ranging from 0.00 to 0.20 with increments of 0.04

other image matching domains. Precision is the ratio of relevant images among the retrieved images whereas Recall is the ratio of relevant images that have been retrieved among total relevant images. The Precision and Recall are given by Equations 4.17 and 4.18, respectively.

$$\text{Precision} = \frac{\text{True} \quad \text{Positives}}{\text{True} \quad \text{Positives} + \text{False} \quad \text{Positives}} \tag{4.17}$$

$$\text{Recall} = \frac{\text{True} \quad \text{Positives}}{\text{True} \quad \text{Positives} + \text{False} \quad \text{Negatives}} \tag{4.18}$$

For computing the Precision-Recall curve, each image in the dataset is taken as a sensed image and compared with all other images in the database to compute the matching score. The matching score is calculated by computing the Euclidean Distance using Equation 4.16 between the Variance Vectors. As discussed in Section 4.4, the lower the distance between images, the higher the similarity between images and vice versa. For computing the Precision-Recall, the images are arranged in descending order based on matching score and then Precision-Recall is computed by using Equations 4.17 and 4.18, respectively.

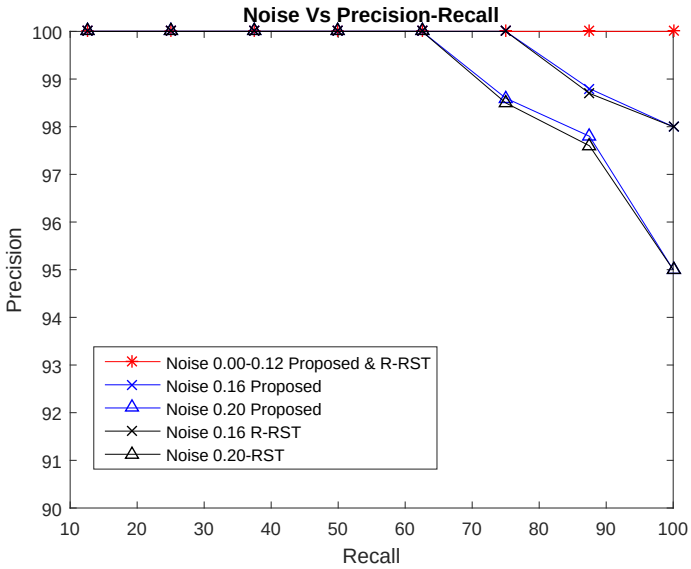


Figure 4.8: Precision-Recall curves of the proposed and R-RST method at different noise densities 0.00:0.04:0.20

Figure 4.8 shows the Precision-Recall curves of R-RST and the proposed method. The results shows that both methods have ideal performance for the noiseless dataset, however, the performance for R-RST and the proposed method slightly decreases as the noise in the images increases. 100% precision is achieved even at a noise density of 0.12. However, at higher noise densities, the precision starts decreasing as the recall starts increasing. The minimum precision achieved at a noise density of 0.20 is 95% which is still higher than the existing region based descriptors such as Hu moments [202], Radon composite features (RCF) [203], Generic Fourier descriptors [204] and Zernike moments [205]. The performance of the proposed approach is as accurate as the performance of the R-RST method. Hence, the proposed method is able to achieve the best Precision in less computation time as compared to the R-RST method.

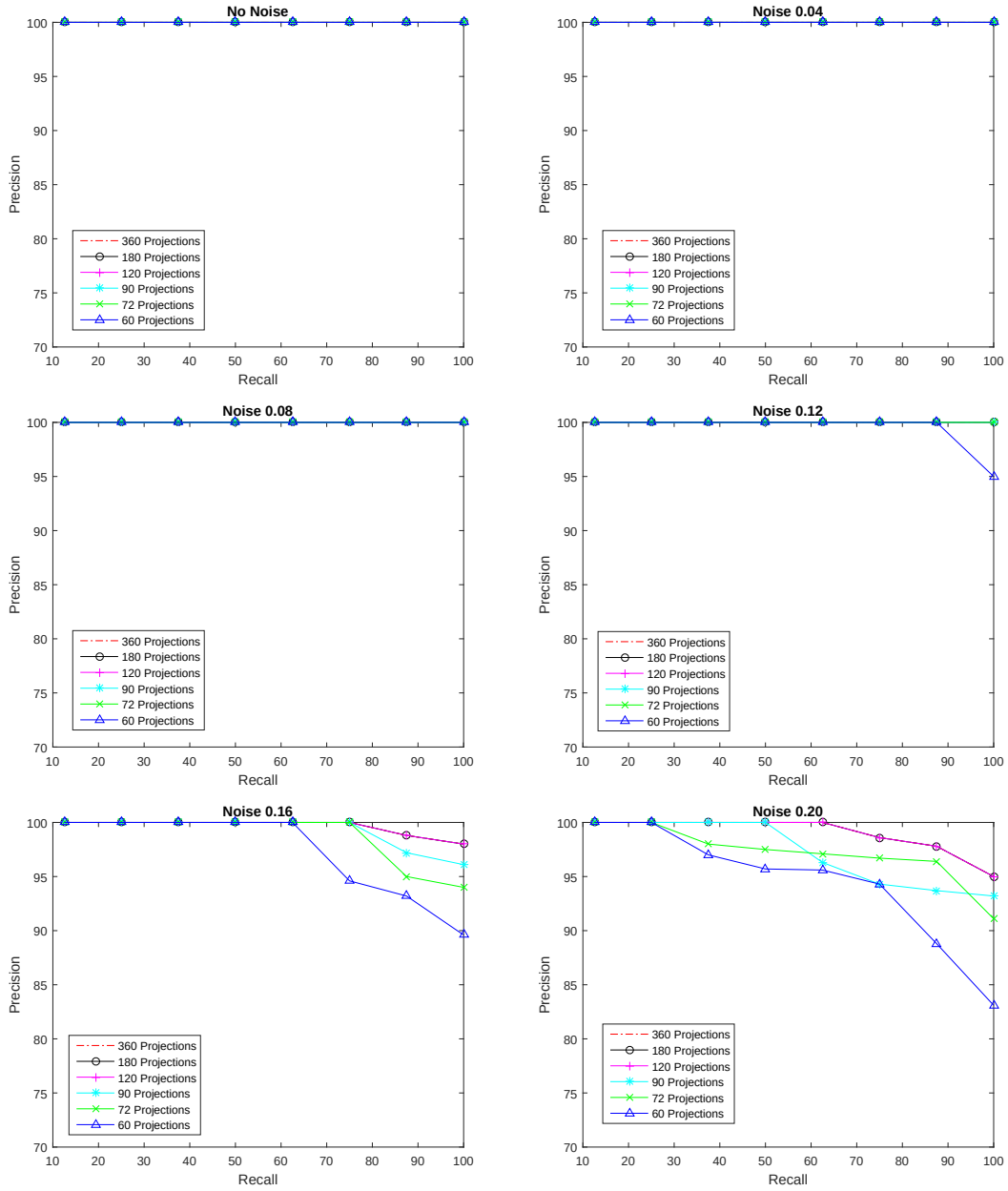


Figure 4.9: Precision-Recall of the proposed method at different noise densities 0:0.04:0.20 and with different numbers of projections

4.6.1 Effect of the Number of Projections on Precision-Recall

Although the proposed method has reduced the computation time of the matching process by using the rotation invariant Variance Vector, the computation time can be further improved. One of the factors on which the computation time of Radon transform depends is the number of projections from 1 to 360 degrees. The higher the number of projections, the larger the computation time and vice versa. The proposed matching process is independent of rotation angle and uses the rotation invariant Variance Vector for matching images which can be computed by using fewer projections. Using fewer projections improves the computation time of the matching process. As the number of projections are decreased, the amount of captured distinguishing information is decreased, which in turn can affect the precision of the matching process. Therefore, the effect of the number of projections used for computing the Variance Vector is studied as Precision-Recall behaviour. However, the number of projections cannot be decreased in [14], because this method matches the images by computing the parameters. If the number of projections are decreased then the rotation angle computation will result in a wrong value, which will affect the precision.

To do this study, the same experiment of computing the Precision-Recall as discussed in the previous section is performed, except the number of projections used for computing the Radon transform is decreased. Figure 4.9 shows the effect of the number of projections on Precision-Recall. The experiments are done on the same logo dataset with noise and without noise. The salt and pepper noise is added to the images with noise density varying from 0 to 0.20 with increments of 0.04. The experiments are conducted for 360, 180, 120, 90, 72 and 60 projections. Projections are sampled regularly at every n projections for 1 : n : 6.

For the noise density from 0 to 0.12, the precision is not affected at all even when the number of projections are reduced to 1/6 of the original number of projections. In other words, the same precision is achieved for both 60 and 360 projections, but the computation time of the Radon transform is decreased to 1/6th of its original time. For a noise density of 0.12, the precision drops to 95% at recall 100. At a higher noise density of 0.16, the same precision is achieved by 120 projections as is achieved by 360 projections, reducing the computation time to 1/3 of the original time. However, the precision starts decreasing as the number of projections decreases further. Nevertheless, we can still achieve a precision of 89.6% by using 60 projections which is better than the performance of other descriptors such as Hu moments [202], Radon composite features (RCF)[203], Generic Fourier Descriptors (GFD) [204], Zernike moments [205] and wavelet ϕ R-signature (Wϕ R)[163]. As stated in the paper [163] the R-RST descriptor performs best, as compared to all the other descriptors, with a precision of 95% at recall 100% at noise density 0.20. The proposed method can achieve, a precision of 95% even with 120 projections at the noise density of 0.20 which is 1/3 of the projections required by R-RST method.

The results show that noisy images require more projections to achieve the same precision as compared to the noiseless images. In other words, the number of projections required depends on the quality of images. The higher the quality of the image, the lower the number of projections required for matching and vice versa. This is because for noiseless images, a lesser number of projections can capture enough distinguishing information, whereas for noisy images, a greater number of projections are required to capture enough distinguishing information.

On the other hand, improving the computation time by reducing the number of projections is not possible in the R-RST method, as the matching process relies on the accuracy of the rotation angle computed between two images. The accuracy of rotation angle computation depends on the number of projections taken

[14]. If all the Radon projections from 1 to 360 degrees are not taken, then the computed rotation angle will not be accurate. This affects the accuracy of the computed translation value between two images which is crucial for alignment. Hence, this results in the misalignment of images and consequently results in the mismatching of images. However, the proposed method is independent of the rotation angle, therefore, it is feasible to reduce the number of projections to improve the computation time while achieving the same precision.

4.7 Conclusion

In this chapter, we have developed an alignment-free matching method based on the Radon transform, for images having pixel-to-pixel correspondence. The alignment-free approach of matching images has eliminated the need for computing the rotation difference between images, which in turn reduces the computation time of the proposed matching method. The effectiveness of the method is tested for object recognition. The results are compared with the existing approaches. The experiments are done on the UMD logo dataset with or without adding noise. The retrieval scores on the UMD logo dataset shows superior performance as compared to well known descriptors for both noisy and noiseless datasets. The results demonstrate that the proposed method is not only computationally faster and accurate but also robust up to 20% of the noise. Moreover, the computation time of the proposed approach is further improved by reducing the number of Radon projections by maintaining the precision which is not possible in existing approaches. Hence, as compared to existing techniques, the proposed method does not require first computing parameters and then matching, but can directly match. In addition, the proposed approach can also match images when only Radon transforms of images are given and no pixel based information is given.

In Chapters 3 and 4, the methods of aligning and matching images having pixel-to-pixel correspondence is discussed. Whereas, in Chapters 5 and 6, the method of aligning and matching partial images (having no pixel-to-pixel correspondence) is discussed.

Chapter 5

Similarity Transformation Parameters Recovery in Partial Images

5.1 Preamble

Chapter 3 discussed a method of computing the parameters relating two images. The proposed method is able to find the parameters more accurately and efficiently as compared to the existing techniques. Moreover, the robustness of method is verified against noise. Results are reported for test images of varying classes and the performance is compared with existing techniques. Although the proposed method is accurate and efficient, the method is applicable only to full images. In other words, the method is appropriate for the images having exact pixel-to-pixel correspondence between them. This is because, in full images, it is possible to find scale by using pixel values, compute centroids of images, and apply a rotation angle estimation method. However, it is not guaranteed that images have exact pixel-to-pixel correspondence all the time. Two such examples of partial images in different applications are biometric identification and medical image diagno-

sis. Each sensed partial image might not have exact pixel-to-pixel correspondence with the reference partial image. However a smaller parts of the sensed partial image may have exact pixel-to-pixel correspondence with smaller parts of the reference partial image. Therefore, due to the unavailability of exact pixel-to-pixel correspondence for partial images, the method proposed in Chapter 3 cannot be used for computing parameters between them. Hence, we need a method that can compute parameters accurately and efficiently between partial images as well.

Partial images play a crucial role in computer vision and image processing applications. This is because it is not always possible to have full images in different applications such as fingerprint recognition and medical image diagnosis. For example, fingerprints captured by sensors of smart phones or fingerprints found at crime scenes are generally partial. Different impressions of the same finger are never identical due to elasticity of skin, noise or other factors [3]. Also, in medical images, the sensed image can be obstructed by some other object, resulting in only part of the sensed image having content in common with the reference image. Moreover, the acquisition of images not only results in partial images, but also images that contain regions of low quality.

From the discussion, it is clear that many applications encounter the problem of partial images. For computing the parameters between partial images accurately, we need a method that can use the common information present them. In this chapter, we propose a region based method to compute the parameters between partial images lacking exact pixel-to-pixel correspondence (shown in Figure 5.1). The region based approach uses the rotation invariant Variance Vector defined in Chapter 4. As it is difficult to find accurate centroid locations in partial images, a method to find a common reference point between images is proposed. Finally, the method of computing the angle is applied to regions of images around the common location instead of applying on entire image. Interestingly, the method of finding common location is applicable to both partial and full images.

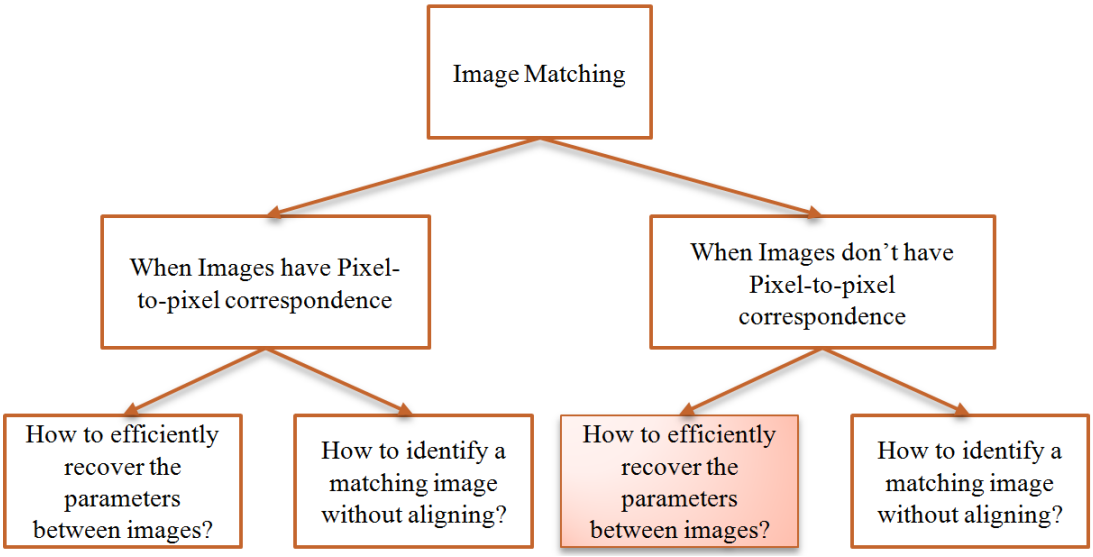


Figure 5.1: The coloured block represents the thesis objective addressed in this chapter.

This chapter is organised as follows: First, the limitations of partial images are discussed in Section 5.2. The issues are followed by requirements for computing the parameters in partial images in Section 5.3. In Section 5.4, an appropriate feature for the computation of parameters in partial images is discussed. Furthermore, the method of extracting a common region between two images is discussed in Section 5.5. The optimization technique of common region identification is introduced in Section 5.5.1. Subsequently, the method of computing the parameters is given in Section 5.6. Finally, the conclusion of the chapter is discussed Section 5.7.

5.2 Partial Image Limitations

Despite the advances in parameter computing methods for full images, finding the parameters of partial images still suffers from a number of problems. In existing techniques of parameter computation, it is assumed that images are of the same size and include the complete object. However, this assumption is not valid in case of partial images. Often, the sensed image differs from the ones present

in database and can be partial in nature. Partial images could be generated in different scenarios.

i) A sensed image can be damaged or distorted and consequently only a small part of the captured image is valid.

ii) A sensed image might have different parts of varying quality. The poor quality regions may not be useful for processing. Hence, only a small recognizable part of the image is available.

iii) A part of the sensed image is obstructed by other object.

iv) In the case of fingerprints, the sensed images obtained from crime scenes or images captured by smart phone sensors are often unclear and partial [209].

v) Due to the miniaturization of sensors[1], the overlap between a sensor and a finger is small and therefore, the resultant sensed fingerprint is partial and can be of poor quality.

The accurate alignment of partial images is mandatory in many applications. The forensic applications need to identify the fingerprints found at the crime scene. The fingerprints found are usually partial and unclear. That makes the matching of these fingerprints difficult. Manual matching of these fingerprints requires the correct alignment of images. The misalignment of images can result in matching of crime scene fingerprints to an innocent person's fingerprint. Another example is medical image registration that demands the correct alignment of images as clinicians require aligned images to make decisions.

The partial images have issues that are superficially similar to full images but also have some of their unique difficulties. These image difficulties challenge the use of existing approaches of alignment. Conventionally, the images are aligned by

[1]The invention of small chip based sensors has made fingerprint identification useful for commercial applications as well [127].

using the extracted features and the accurate extraction of features is a difficult task. The extraction of features require image processing which is time consuming. This method of aligning the images poses more challenges in the case of partial images, as a partial image does not consist of enough features. Having varying quality in different parts of image makes it difficult to extract the features. Another way of aligning the images is by computing the parameters as discussed in Chapter 3. However, the parameter computation using different transforms is difficult in partial, distorted and unclear images, because, these methods need pixel-to-pixel correspondence between images. Partial images lack this correspondence.

The method proposed in Chapter 3 computes the parameters between images using the Radon transform. However, the approach for finding the parameters encounters problems in the case of partial images because of two main reasons:

i) The method requires finding the centroid of images for computing translation and rotation angles. However, it is difficult to find the exact location of the centroid in partial images as different parts of the object may be present in images. Therefore, if this centroid location is used, it will result in an error in translation and rotation angle computation.

ii) To find the rotation angle, the Row Variance Vector is computed by using the Radon transform. However, this method is applicable only to images having exact pixel-to-pixel correspondence which is possible only in full images. When images are partial, only a part of object is present in each image, hence, images do not have exact pixel-to-pixel correspondence. This will result in different Radon transform for both images. Due to different Radon transforms, the computed Row Variance Vectors of both images will be different that will compute the wrong value of rotation angle.

Therefore, a method proposed in Chapter 3 cannot be used for partial images. As discussed above, partial images have many limitations that need to be taken

care of. A method for computing parameters requires the ability to make use of whatever information exists in a partial image. In the following section, we discuss the requirements of a parameter computation method for partial images.

5.3 Prerequisite for Computing Parameters

The method discussed in Chapter 3 for computing the parameters has a basic requirement that images should have exact pixel-to-pixel correspondence. However, in partial images only parts of images have pixel-to-pixel correspondence. Therefore, for computing the parameters between partial images, a common part between the images needs to be identified. To use the method proposed in Chapter 3 for computing parameters, we need a method that is able to extract the common part between images which exhibits pixel-to-pixel correspondence. The extracted common part fulfils the condition of method proposed in Chapter 3, therefore, the common regions between the images can be used to compute the parameters.

The other requirement of the method proposed in Chapter 3 is the use of centroids of images as indicators of a common location. The centroid of images cannot be used as a common location in partial images, because it might not be possible to detect the correct location of a centroid when the images are partial. In the case of full images, the images coincide perfectly when aligned by using the centroids as indicating a common location, however, partial images will not coincide if aligned by using centroids as indicating a common location. This is because, partial images consist of different parts of the same object and have only a part in common. Therefore, using the centroids of partial images as indicators of a common location is not an appropriate choice. Hence, we need to find a common location between both images before applying the method proposed in chapter 3 for computing parameters.

5.4 Region as a Feature

The major problem in designing a parameter computation method for partial images is the selection of appropriate information. One possibility is to use feature points of an image forcomputing the parameters. However, partial images due to their incomplete nature do not possess the requisite number of features that would support parameter computation. In other words, partial images have less discriminative information as compared to full images. In addition, there is a high probability of misidentifying the features in poor quality or distorted images. Moreover, using only one kind of feature for partial images might not be enough to accurately compute the parameters. Therefore, we need to incorporate more distinguishing information instead of only relying on a few image features such as corner or edge points.

In [3] and [194], it is also stated that additional features increase the discriminating information that can be used to increase performance of the system. In order to address these problems, we need to make use of all features present in images [10]. In [10] and [179], authors have stated that gray-level information from pixels around feature points contain richer information about the local region than attributes of the feature points. Moreover, the information used for matching images should be able to distinguish the inter and intra images. Therefore, computing the parameters of images by using the pixel based information is an optimistic approach.

The pixel based information of images can be captured by using regions of images. Regions of images represent all dimensional attributes such as microscopic features like corner, tee points, minutiae, pores etc. and macroscopic features like singularities, blobs etc. As all the pixels contribute, the pixel based information is less sensitive to quality as compared to specific features of images and minimizes the effect of an insufficient number of features, error in feature extraction and ab-

sence of features. Moreover, the region of images represents only local information of an image, therefore, region based techniques can deal with non-linear distortion present in images. Hence, regions of images can be used for computing the parameters between partial images.

5.4.1 Shape of Region

As discussed in the section above that it is required to extract a common region between images for computing the parameters. However, as images are unaligned, finding a common region between images in pixel domain is difficult. The major cause of misalignment of images is the translation and rotation differences between them. In a region based approach, the translation difference of images can be dealt easily by using a sliding window approach. However, dealing with the rotation difference between images is time consuming. In [10], the region based technique is used for extracting the common regions between fingerprints. For that, the approach computes correlation between cropped square shaped regions. However, for computing correlation, regions need to have pixel-to-pixel correspondence between them, whereas, square shaped regions of unaligned images do not have pixel-to-pixel correspondence. Therefore, for computing the correlation between rotationally unaligned regions of images, one of the regions need to be rotated for each possible angle as compared to other regions and correlations is computed for all angles. The angle that results in the maximum correlation value is considered as the rotational angle between images.

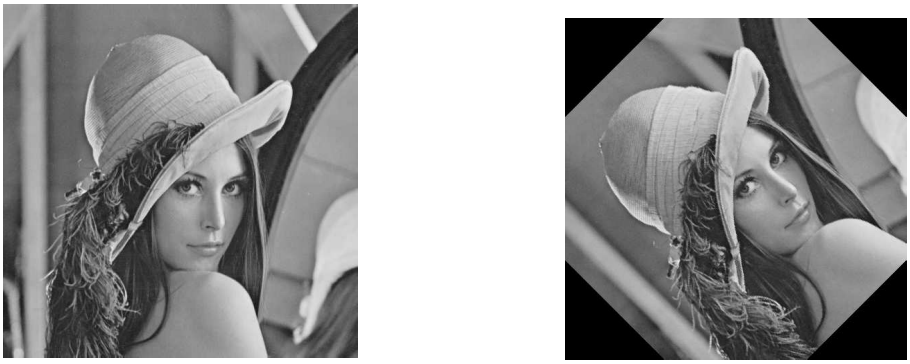
Although this process results in finding the correct matching region, it is a time consuming process. Therefore, to eliminate the exhaustive process of rotating one region with respect to another to find a common region, regions can be converted into other domains such as the Radon domain to make it easier to deal with the rotational difference. As discussed in Chapter 3, the Radon transform simplifies

the rotational differences between images by converting the rotation in the pixel domain to a shift in the Radon domain. Hence, the translational and rotational difference of images can be processed easily.

The rotation property of the Radon transform (discussed in Section 4.3) is applicable only when the two images have pixel-to-pixel correspondence except being rotationally unaligned. As we are using the regions of images, the shapes of regions play an important role. In other words, regions should be of a particular shape such that they possess pixel-to-pixel correspondence even when images are unaligned. If the reference and sensed images' regions do not have pixel-to-pixel correspondence then the computed Radon transforms are different and cannot be applied to find common regions.

Figures 5.3 (a) and (b) show square shaped reference and rotated image regions cropped from the reference and rotated Lena images (shown in Figure 5.2). The regions are not only rotationally unaligned but also do not have exact pixel-to-pixel correspondence. To verify the pixel-to-pixel correspondence between regions, the initially rotated image region is rotated back by the same angle (as shown in Figure 5.3(c)). It can be seen that the rotated and reference images' regions do not have pixel-to-pixel correspondence, because the pixels present in the corners of both regions are different. In other words, the rotation difference does not allow such square shaped rotationally unaligned regions to have pixel-to-pixel correspondence. Hence, regions should be of such a shape that they have pixel-to-pixel correspondence even when images are rotationally unaligned.

A circular shape results in pixel-to-pixel correspondence between regions even if images are rotationally unaligned. Figures 5.4 (a) and (b) show reference and rotated circular shaped regions cropped from reference and rotated Lena images shown in Figure 5.2. To check the pixel-to-pixel correspondence, the rotated image region is rotated back by same angle by which Lena image has been rotated initially. It shows that the rotated and reference image regions have exact pixel-to-pixel



(a) (b)

Figure 5.2: Lena reference and sensed (45° rotated) images

(a) (b) (c)

Figure 5.3: (a) Square shaped Lena reference image region; (b) Square shaped rotated (45°) Lena image region; (c) Sensed image region rotated back by 45°

(a) (b) (c)

Figure 5.4: (a) Circular shaped Lena reference image region; (b) Circular shaped rotated (45°) Lena image region; (c) Sensed image region rotated back by 45°

correspondence except being rotationally unaligned. This fulfils the condition of Radon transform, hence, rotationally unaligned circular regions result in identical Radon transforms. Therefore, for partial images, circular regions can be cropped from images and used for finding the common region between both images.

5.5 Method of Common Region Identification

As we have stated, in partial unaligned images a common region between images needs to be extracted for computing the parameters. Generally, the misalignment between images exists because of translation and rotation differences. Therefore, to find a common region between images we need to take care of both differences between them. To avoid the translation effect for finding the common regions, the naive approach of a sliding window is used. In this approach, a region cropped from the sensed image is compared with all possible regions of the reference image thereby cropping a region around each pixel of the reference image.

For dealing with the rotation difference, the Radon transform of regions can be used. The rotation property of Radon transform states that rotation in the pixel domain is a shift in the Radon domain. Therefore, the distance between the Radon transforms of regions can be computed and the region that results in a minimum distance can be considered as a matching common region. This is because, two identical but rotationally unaligned regions result in identical Radon transforms, except being circularly shifted as shown in Figure 3.3. Hence, for computing the distance between Radon transforms, one Radon transform needs to be shifted for each angle between 0 to 2π as compared to the other transform. Although, this method can result in finding the correct matching region, it is computationally exhaustive like a correlation method as Radon transforms are shifted for each angle and then distance is computed.

Therefore, for efficient common region extraction, the proposed rotation invariant Variance Vector (discussed in Section 4.3) is used. As discussed, comparing two regions by using correlation or the Radon transform is a time consuming process, so regions can be converted into rotation invariant representation and compared, without worrying about rotation differences between images. The procedure of

finding the common regions using the rotation invariant Variance Vector is given below:

i) Compute the Variance Vectors for all possible regions of a reference image (required for the sliding window approach).

ii) Crop a region from the sensed image and compute its rotation invariant Variance Vector.

iii) Compare the Variance Vector of a sensed image region with the Variance Vector of each reference image region by computing the Euclidean Distance.

iv) Choose a reference image region whose Variance Vector results in a minimum distance with respect to the Variance Vector sensed image regions.

The rotation property of a Variance Vector states that the Variance Vectors for two rotationally unaligned but identical images are identical. In other words, rotationally unaligned regions having pixel-to-pixel correspondence results in identical vectors and therefore, the Euclidean distance between vectors is zero. However, due to noise present in images, two regions are never identical that cause variations in their Variance Vectors and hence, the Euclidean Distance between vectors result in a non-zero value. Therefore, a region from the reference image is selected whose vector results in a minimum distance with the sensed image region vector. That region is taken as a matching region. The complete process of finding the common regions by using the Variance Vectors is shown in Figure 5.5. The regions found by Variance Vectors are verified by comparing with the common regions found by a correlation method. Hence, the Variance Vector helps in efficiently identifying the common regions between the images without aligning them. The algorithm for finding the common region is given in Algorithm 5.1.

Algorithm 5.1 Find a Common Matching Region between Partial Images.

Step 1: Compute Variance Vectors of the Reference Image Regions:
$Col \leftarrow$ Number of columns in reference image
$Row \leftarrow$ Number of rows in reference image
$Cnt \leftarrow 1$ (Counter for the number of reference image regions)
for $i = 1$ *to* Col **do**
  for $j = 1$ *to* Row **do**
    Crop a region at location (i,j) of diameter D
    Compute the Variance Vector of a cropped region ($V_R(1 : D, Cnt)$)
    Cnt=Cnt+1;
  end for
end for

Step 2: Compute the Variance Vector of a Sensed Image Region:
Crop a region of diameter D from the sensed image
Compute Variance Vector of cropped region (V_S)

Step 3: Compute the Euclidean Distance (ED) between Reference and Sensed Image Regions:
for $K = 1$ *to* $Cnt - 1$ **do**
  ED(Cnt) $= \sqrt{\sum_{u=1}^{D}(V_S(u) - V_R(u, Cnt))^2}$, where D is the diameter of circle
end for

Step 4: Identify a Common Matching Region:
Select a reference image region that results in a minimum ED with a sensed image region as a matching common region.

5.5.1 Optimisation of Common Region Extraction Method

As discussed in Section 5.5, the common regions are identified by using a sliding window approach which is a time consuming process. Therefore, for improving the efficiency of a common region extraction approach, a few optimization techniques can be applied:

As discussed earlier, common regions can be identified by either using Radon transforms or Variance Vectors. However, if we use the Radon transforms for identifying common regions then for each region comparison, the number of computation steps required is $P^2 \times D$ where, P is the number of projections between 0 to 2π and D is the diameter of circular region. On the other hand, if Variance Vectors are used then the number of computation steps for each region compar-

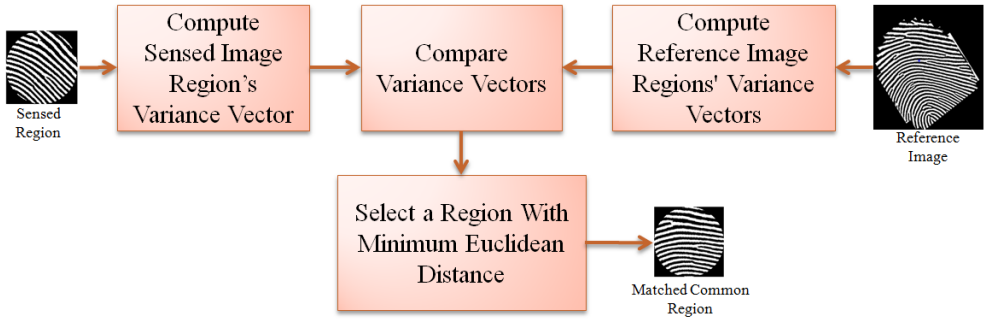


Figure 5.5: The process of identifying common regions between unaligned partial images without aligning them

ison is D. Hence, by using a Variance Vector we have decreased the number of computation steps for each comparison by P^2. This makes a significant difference in computation time when the number of regions in the sensed image is large. If the number of regions in each image is Cnt then by using Variance Vector we have reduced he number of computation steps by $Cnt \times P^2$.

The Variance Vectors of sensed image regions are compared with all regions of reference image regions to identify the common region. Therefore, to make the identification process efficient, the Variance Vectors of the reference image regions can be computed off-line and stored in a database. This approach converts an exhaustive regions comparison process (which include vector computation and comparison) to a simple vector comparison process.

Furthermore, as shown in Figure 5.6, a Variance Vector is a symmetric vector. Therefore, instead of comparing all the components of vectors, only half the components need be compared. This process further reduces the number of computation steps for each comparison from D to $\frac{D}{2}$. Moreover, when the Variance Vector of a circular region is computed, the vector consists of zeros at the beginning and end of a vector due to the zero pixels in the corners of regions (as shown in Figure 5.4). This increases the number of computations unnecessarily, hence, the leading and trailing zero valued components of the vectors can be trimmed. The remnants of the vectors are compared to decrease the computation steps.

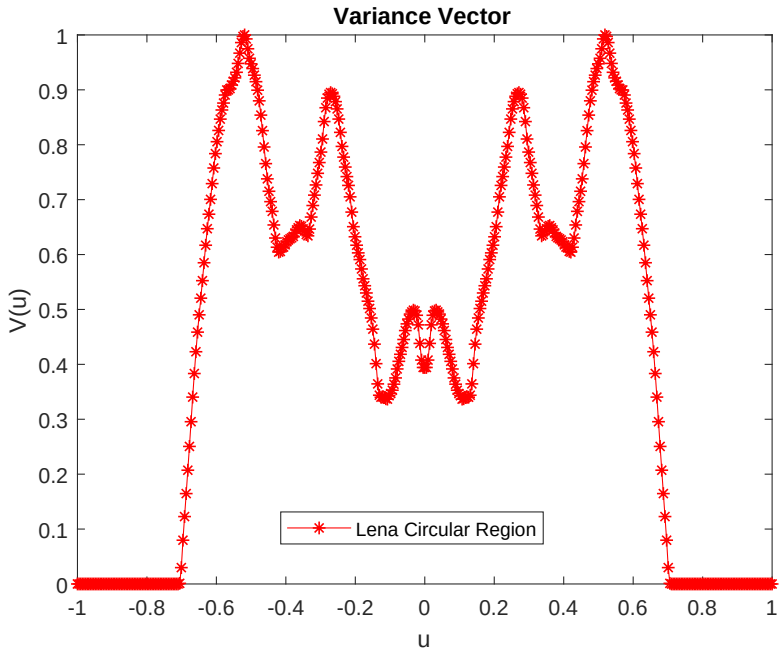


Figure 5.6: The Variance Vector of the Lena image circular region shown in Figure 5.4 (a)

Hence, by using the above described techniques, common regions between partial images can be extracted efficiently without affecting the accuracy.

5.6 Method of Computing Parameters using a Common Region

As discussed in Chapter 2, alignment of partial images is very important for several image processing and computer vision applications. One of the important applications where alignment of images has no pixel-to-pixel correspondence is medical image registration in which images can be obstructed by the presence of other objects. For alignment, we need to compute the transformation parameters relating to images. Figure 5.7 shows reference and sensed partial Lena images lacking pixel-to-pixel correspondence. The sensed image has rotational difference

ϕ_0 and translational difference (x_0, y_0). The idea is to compute both rotation and translation parameters between partial images.

Figure 5.7: (i) Lena reference Image (I_R); (ii) Lena sensed image (I_S) obtained by applying the rotation and translation transformations [6, 4, 4]

In Chapter 3, we have proposed a method for computing parameters. To apply this method for computing parameters, there are two requirements as discussed in Section 5.3: (i) images should have pixel-to-pixel correspondence; and (ii) a common location is required between images. Although partial images do not have pixel-to-pixel correspondence, a common region having this correspondence can be extracted by using the method proposed in Section 5.5. Once the common regions are extracted, the next issue is to find a common location between regions. For that, the centroid of extracted common regions between partial images that have pixel-to-pixel correspondence can be used as a common location between them. As circular regions are used for finding the common regions, the centre of a circle can be used as a common location (like the centroid used in full images) because the centre and centroid of a circle exist at the same location. Hence, for a reference image region, the location of the centroid is represented by C_{Rx} and C_{Ry}; for a sensed image region by C_{Sx} and C_{Sy}.

Once the common locations are chosen, the next step is to find a rotation angle between images. For computing the the rotation angle, the Radon transform

of common regions of unaligned images are computed. The Radon transforms of regions cropped from reference (I_R) and sensed (I_S) images are represented by $RT_{Rr}(\phi, u)$ and $R_{Sr}(\phi, u)$, respectively. As we know that the Radon transform shifts the rotation in the pixel domain to the shift in the Radon domain. Hence, this property is used for computing the angle by comparing the Radon transform for each possible angle. However, comparing the Radon transform for each angle is a computationally expensive. Therefore, to increase the efficiency of the computation process, instead of comparing the Radon transforms, we can simply compare the Row Variance Vectors (discussed in Section 3.4) of both regions. The Row Variance Vectors of reference and sensed image regions are represented by RVV_{Rr} and RVV_{SR}, respectively. The rotation angle between images is computed by using the objective function:

$$H(\phi_0) = \int_0^{2\pi} |RVV_{Rr}(\phi) - RVV_{Sr}(\phi - \phi_0)| d\phi \tag{5.1}$$

Where RVV_{Rr} represents the Row Variance Vector of the reference image region and RVV_{Sr} represents the Row Variance Vector of the sensed image region. The rotation angle ϕ_0^* between angle 0 to 2π is given by:

$$\phi_0^* = arg \quad min_{\phi_0}(H) \tag{5.2}$$

The next parameter to compute is translation between two images. To find the translation, we already know the rotation angle (ϕ_0) between images and the reference points locations at (C_{Rx}, C_{Ry}) and (C_{Sx}, C_{Sy}) in the reference and sensed images, respectively. The translation parameters (x_0, y_0) can be computed by Equation 5.3 and Equation 5.4, respectively.

$$y_0 = C_{Sy} + C_{Rx} \sin(\phi_0^*) - C_{Ry} \cos(\phi_0^*) \tag{5.3}$$

$$x_0 = C_{Sx} - C_{Rx}\cos(\phi_0^*) + C_{Ry}\sin(\phi_0^*) \tag{5.4}$$

The step by step procedure for computing transformation parameters is:

i) Identify a common region between partial images by using Algorithm 5.1.

ii) Take the centre of common regions as common locations between the reference image (C_{Rx}, C_{Ry}) and the sensed image (C_{Sx}, C_{Sy}).

iii) By using Equation 3.3, compute Radon transforms of the reference and sensed images common regions RT_{Rr} and RT_{Sr}, respectively .

iv) By using Equation 3.21, compute Row Variance Vectors of the reference and sensed images' regions (RVV_{Rr} and RVV_{Sr}), respectively.

v) Compute the objective function using Equation 5.1.

vi) Deduce the rotation angle using Equation 5.2.

vii) Compute the translation coordinates using Equations 5.3 and 5.4.

Hence, the proposed method can be used for computing transformation parameters between partial images where most of the existing methods fail.

5.7 Conclusion

In this chapter, a strategy to compute parameters in partial images having no pixel-to-pixel correspondence has been presented. As the partial images lack in features, it is proposed that regions of images can be used as features in such images. Also, the shapes of regions play an important role in computing parameters, hence an appropriate shape of region is described. The major issue in partial images of not having exact pixel-to-pixel correspondence is dealt by presenting a method of extracting a common region between images having this correspondence. Moreover, the optimization techniques for extracting the common regions

efficiently is provided. Finally, a region based method is proposed that can compute the parameters of partial images where most of the existing methods fail.

In this chapter, we have provided the theory of dealing with the issues of partial images. In Chapter 6, the proposed theory is tested for identifying partial images and the approach is tested on standard fingerprint datasets.

Chapter 6

Region-based Alignment-free Partial Image Matching

6.1 Preamble

We have discussed an alignment-free method of matching images having pixel-to-pixel correspondence in Chapter 4. The method not only performs better in terms of accuracy but is also efficient in matching images. However, the method of computing similarity and matching images is applicable only when images have pixel-to-pixel correspondence. Therefore, the approach is not applicable to partial images as partial images lack this correspondence. However, parts of partial images may have the required correspondence. In that case, for matching partial images, it is essential to extract the common regions between images and these extracted regions can be used for matching the partial images. Due to these problems, a region based tranformational approach is proposed to match partial images without aligning them. The proposed approach in this chapter (shown in Figure 6.1) exploits property of the Radon transform and uses a sliding window approach to match the images. The chapter presents an alignment-free region based image identification technique that is applicable to all types of images.

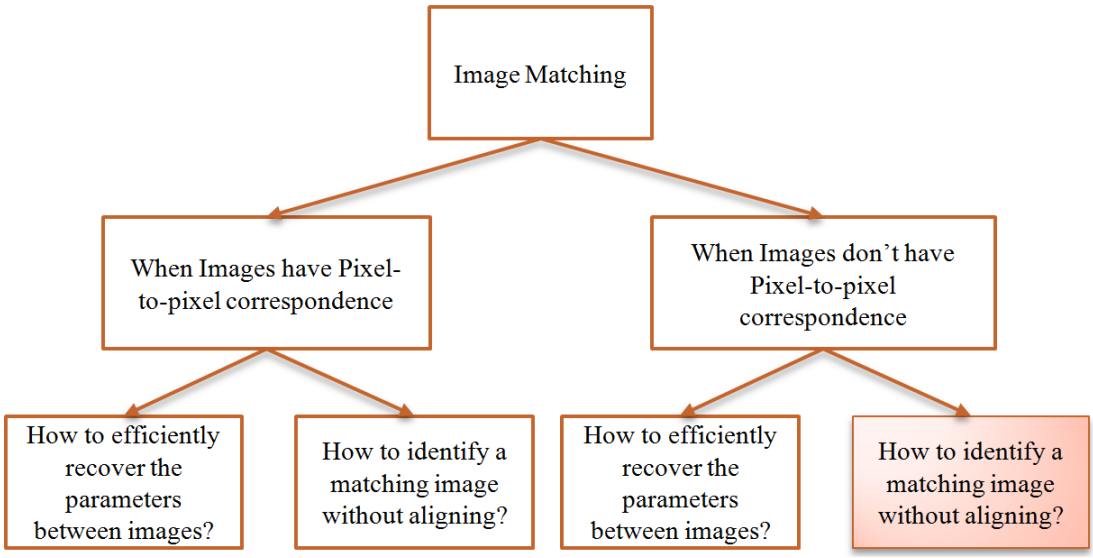


Figure 6.1: The coloured block represents the thesis objective addressed in this chapter.

The rest of the chapter is organised as follows. Beginning with the issues, the difficulty of matching partial images is explained in Section 6.2. The requirements for developing a method for matching partial images follows in Section 6.3. Section 6.4 explains how to meet the requirements of matching partial images by describing: (i) appropriate features for matching partial images; (ii) the extraction of common regions between images; and (iii) a new similarity metric that is independent of any image specific features and applicable to partial images. Section 6.5 discusses the performance of the proposed approach by using a fingerprint dataset as an example in comparison with existing techniques. This section also discusses a method to deploy the maximum possible available information in partial images by using multiple regions. It also studies the effect of the number of regions, size of regions, noise and number of Radon projections on accuracy. This section explains the computational complexity of the proposed approach as well. Evaluation of results show the superiority and appropriateness of the proposed approach. Experimental results are followed by a conclusion drawn in Section 6.6.

6.2 Issues in Partial Image Matching

Many automated image identification systems have been successfully implemented and it has widely been used in different commercial applications. [58] states that the exquisite implementation of image identification manifested in different applications shows the existing technology to be robust. However, that is not true and there are still many issues that need to be resolved. One of the major issues is in the identification of partial images. In the research community a number of research projects are still in progress to improve the accuracy of identification of partial images. However, the existence of a fully automated partial image identification system is not known yet [16].

Despite the advances in image identification systems, matching of partial images still suffers a number of problems. The matching techniques for full images are quite mature as compared to partial image matching [10]. Usually in existing matching techniques it is assumed that images are of same size and include the complete object. However, this assumption is not valid in case of partial images. Generally, the acquired images differ from the ones present in the dataset and can be partial in nature. Different types of partial fingerprints can be captured from different scenarios as shown in Figure 6.2. The matching of these partial fingerprints is a crucial issue that needs to be addressed. One day-to-day life example is the forensic investigation to identify the fingerprints found at a crime scene. The FBI maintains ten prints (impressions of all ten fingers) of all well known criminals. It is time consuming to find the suspect from big datasets. Therefore, it would be of enormous use in different applications if partial images could be matched accurately and efficiently.

The identification of partial images has issues that are superficially similar to full images, but also has some unique challenges such as, the non-availability of enough distinguishing information and different qualities in different parts of im-

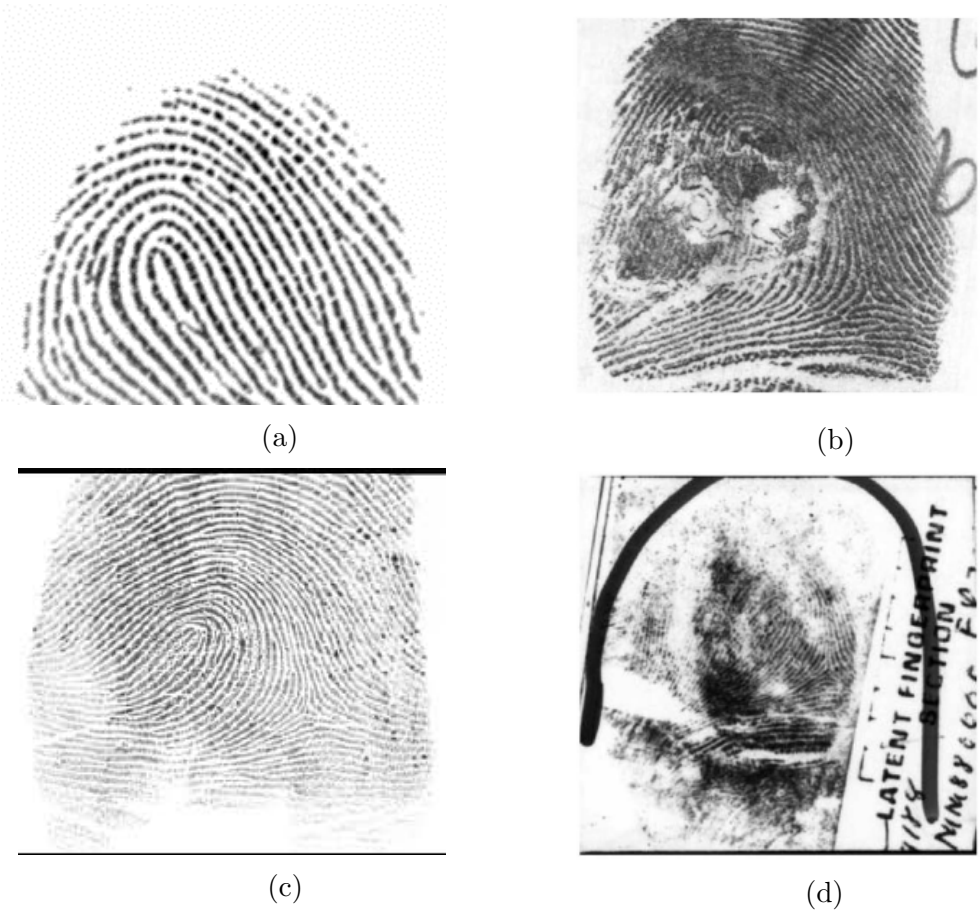


Figure 6.2: Different scenarios that result in partial fingerprints (a) miniaturization of sensors; (b) a distorted fingerprint; (c) a low quality fingerprint; (d) fingerprints found at crime scenes [16].

ages. The identification of partial images is complex because partial images can be distorted, small, unclear and noisy. Due to these problems the information contained in partial images is limited. That challenges the use of existing approaches for matching partial images. Some of the challenges in matching partial images are:

i) A partial image does not consist of enough features such as corners, tee points, minutiae, ridge structure and pores. These reduce the distinguishing information and make the matching process difficult.

ii) For computing the global similarity between the images, images need to be aligned first. Generally the images are aligned by using image features.

However, a partial fingerprint may not have required features. That restricts or complicates the alignment process of partial images.

iii) The partial images may consist of background noise that not only hinders the extraction of features but also reduces the amount of distinguishing information.

iv) The low quality of partial images makes the details of images smudgy and unclear, resulting in less common information between images.

v) Images captured in uncontrolled environment (such as fingerprint captured by smart phones) may result in distorted images which may have less distinguishing information [16].

vi) Manual intervention is required at different levels in some partial image matching. For example, in matching of the latent fingerprints, the fingerprints are generally annotated manually and then automated matching system is used to find a list of top k matches ($k > 0$). This list is further verified by the skilled experts to check if individualization exists. This makes the complete process time consuming and laborious.

vii) In partial image matching, one needs to identify which part of object is present in a given image that makes the matching process more difficult.

All the issues discussed above impose a number of challenges in adopting the existing approaches for matching partial image. Next section discusses the requirements for developing a method that can match partial images accurately and efficiently.

6.3 Partial Image Matching Requirements

Most of the conventional techniques of image identification are based on features of images. These techniques have several advantages and disadvantages as shown in the Table 2.2. As discussed in Section 2.8, pixel based matching methods are better as compared to feature based methods. However, the existing pixel based methods require pixel-to-pixel correspondence between images which is not necessarily present in partial images. A pixel based technique has been proposed in [10] for images not having pixel-to-pixel correspondence. Although it is a pixel based matching method, it requires alignment of the images and for aligning the images it uses specific features of image. The extraction of those features faces the same challenges as are faced by feature based techniques. Another feature-free alignment based method is also proposed in the same paper but this method uses brute force method for aligning that makes it unsuitable for real time applications.

Therefore, existing techniques are not capable of matching partial images accurately and efficiently. Moreover, in case of many partial image matching approaches manual intervention is required [16]. However, the manual identification is subjected to inconsistencies as well. Moreover, there are not enough number of skilled and certified experts. Hence, there is a need for scientific automated method for matching partial images. However, it is clear that there are a number of requirements that an automated method needs to fulfil for matching partial images. Following is the list of requirements for partial image matching method:

i) The method should be independent of image specific features such as corners, tee points, minutiae, singularities, pores etc. This is because:

   - Partial fingerprints consist of fewer details, as compared to full images, limiting the number of features present in images;

- There is a high probability that partial images contain noise that makes the accurate feature extraction process difficult.

ii) As partial images lack distinguishing information due to limited textural structure, using only fragmentary information (such as only corners or minutiae) may result in false matching. Consequently, for matching, the method should exploit as much of the partial image information as possible.

iii) The method should be tolerant to non-linear distortion as different parts of partial and low quality images consist of different amounts of distortion.

iv) The alignment of images is a crucial step in most of image matching techniques. However, alignment not only adds complexity across the entire matching process, but also inaccurate alignment degrades the accuracy of systems. Moreover, alignment is generally based on image features (such as minutiae or singularity) or parameters need to be computed. However, extracting features and computing parameters are difficult and time consuming processes. Hence, there is a need for a partial image matching method to match partial images without aligning them.

v) As the required method should not only consider all the given information, but also be independent of alignment. A new similarity metric is required which is suitable under these conditions.

vi) Besides fulfilling all the above mentioned conditions, the method should be accurate and efficient.

6.4 An Approach to Partial Image Matching

The unique characteristics of partial images present many challenges in building a highly accurate image matching method. Therefore, to match partial images,

we need a method that can fulfil all the requirements discussed in the Section 6.3. Section 6.4.1 discusses why the regions of images are taken to capture the distinguishing information and how regions can be used to utilize all the available information of partial images. It is also explained how region based information can tolerate non-linear distortion. Section 6.4.2 summarizes the alignment-free technique for finding the common matching region between images. A new similarity metric is proposed in Section 6.4.3 that can be used to compute the global similarity value between images for the given circumstances of partial images.

6.4.1 Features in Partial Images

The major issue in the design of a partial image identification system is the selection of features that can best identify partial images for the underlying matching applications. Partial images due to their incomplete nature do not consist of a sufficient number of features required for matching. The amount of details present in partial images is always less than full images, which means partial images consist of less distinguishing information as compared to full images. In addition, there is a high probability of misidentifying the features in low quality images. As discussed in Section 2.5, using only one kind of feature for identifying partial images might not be enough. Therefore, we need to incorporate more distinguishing information instead of relying only on a few features. [3] and [194] also state that additional features increase the discriminating information that can be used to ameliorate the performance of the system. Therefore, as explained in Chapter 5, the regions of an image can be used for matching partial images, as region incorporates all features present in that region.

Regions of images represent all dimensional attributes such as microscopic and macroscopic features. As all the pixels contribute, pixel information is less sensitive to quality as compared to the specific features of an image and minimizes

the effect of an insufficient number of features, error in feature extraction and the absence of features. Moreover, the region of images represents only local information of the images. Therefore, region based techniques can deal with non-linear distortion present in images. Cappelli et al. [5] mentioned that the local region based matching has addressed the challenges of global matching such as lack of tolerance to non-linear distortion and high computational complexity. As a result, region based techniques can address the issues of matching partial or low quality images.

To improve the performance of a system, most of the available information from partial images is required to be used in matching partial images. One of the possible ways of doing this is by using multiple regions of images. Multiple regions can be of two types: disjoint and overlapping of same size. The same number of disjoint regions, compared to overlapping regions, can cover a larger area of an image and represents more distinguishing information. In other words this implies that, we need to use a larger number of overlapping regions as compared to disjoint regions for capturing the same amount of distinguishing information. The details of disjoint and overlapping regions are given below:

Disjoint Regions: Disjoint regions can be cropped from an image by applying a grid of a size similar to the diameter of a circular region. Although, increasing the number of disjoint regions increases the amount of distinguishing information captured, there are some issues in using disjoint regions:

i) In partial images, there is difficulty in attaining the required number of disjoint regions that can distinguish them from the inter images (images of different objects). Therefore, due to the limited number of disjoint regions, the approach is not suitable for partial image matching.

ii) Extracting the maximum number of disjoint regions having specific features in an image is a time consuming process that can affect the efficiency of

the system. Moreover, if the process of region extraction is made faster but regions are not extracted properly, it is possible that the distinguishing information of the image is left unused. That affects the accuracy of system.

We may take more disjoint regions to improve the accuracy of the system. However, a problem arises when image does not have sufficient numbers of disjoint regions. Therefore, it is necessary to find an alternative method that can make use of most of the available distinguishing information present in images. Such alternative can be overlapping regions and the definition of these regions is given below:

Overlapping Regions: An insufficient number of disjoint regions in partial images prompts the use of overlapping regions in images. The overlapping regions can be cropped in images by using a grid of size smaller than the diameter of a circular region. The use of overlapping regions has many benefits as follows:

i) It can be used for partial images as partial images do not have enough disjoint regions.

ii) It helps in using most of the given information in an image. For example, assume a partial image has only 1 region and the rest of the partial image cannot form another disjoint region. Then, instead of discarding the rest of the information, we can have other overlapping region and use that information. Although, overlapping regions do not increase the amount of distinguishing information in the same ratio as that of disjoint regions, it avoids the inefficient use of given information.

iii) One can expect an increase in accuracy by considering a greater number of overlapping regions. However, it is not possible to have many disjoint regions as an image may not have that many disjoint regions.

iv) As a large number of overlapping regions can be cropped from an image, it can eliminate the effect of low quality regions in the fingerprint. The low

quality regions are likely to have low similarity with the matching region. Therefore, while computing a final similarity score between images, regions with only highest similarity can be considered.

Hence, overlapping regions can offer an opportunity to make better use of the given information of an image. While using multiple regions, the global information of images is used as well. Hence, information achieved at a local level can also be checked for consistency at the global level. However, for finding multiple matching regions, images need to be aligned first. That itself is a time consuming process. The region-based techniques (discussed in [10]) perform best among existing pixel based matching techniques, using both local and global information. Although, this approach achieves the best accuracy, it suffers from time complexity due to its exhaustive alignment strategy. Hence, it will be interesting to investigate the region based techniques further by eliminating the alignment issues. Therefore, the fundamental question that we need to explore: how can region based approach be applied for matching partial images without explicitly aligning the whole or part of the images?

6.4.2 Alignment-free Common Region Identification

As discussed in the previous section, regions of images can be used for matching partial images. However, images are unaligned due to translational and rotational differences between them. Therefore, for matching images without aligning them we need to extract the common regions without aligning them. Such a method of extracting the common region without aligning images is proposed in Section 5.5.

Misalignment of images occurs mainly due to translational and rotational differences between them. The proposed approach for finding the common region deals with both translation and rotation differences. The Radon transform takes care of the rotation difference between images; the sliding window approach can provide

the translation difference between them. Hence, the method can extract common matching regions without aligning the reference and sensed images. Therefore, it eliminates the need for aligning images in determining the common information present in images.

Once common regions are extracted from images, the next step is to compute the similarity between the images. One of the recent region based approaches [10] consists of three steps (i) find common regions between images; (ii) align images and find the corresponding common regions; (iii) compute similarity between images. However, the proposed approach can simply finds corresponding regions between images without aligning them by using the method proposed in Section 5.5 [201]. For computing the similarity, existing measures such as correlation cannot be used as images are still unaligned. Therefore, we have adopted a similarity metric that can compute global and local similarity between unaligned images. The detailed explanation of computing similarity is given in the next section.

6.4.3 Similarity Computation between Unaligned Partial Images

In order to complete an existing matching process, it is required to compute the similarity between images. In partial images, the overlap between images can be quite small. That makes the similarity computation challenging. In existing region based approaches [10], once common regions are found between the sensed and reference images, the next step would be to align images and compute similarity between them. For that, the extracted common region would be considered the reference location between these images, and images would be aligned by using the angle that results in maximum correlation. By using the reference location, corresponding common regions between images are found. The similarity between these local regions is computed and consolidated to get the global matching similarity

score, which is used for identifying the images. As discussed, image alignment would have been necessary for computing the similarity in existing techniques, whereas the proposed method need not align images for computing similarity. This is because, the proposed method can find a common region in the sensed and reference images without aligning them. Therefore, the other corresponding common regions between sensed and reference fingerprints can be found quickly by the same process as well.

After finding the corresponding matching regions between sensed and reference images, the similarity between images can be computed. However, we have converted the image regions into rotation invariant vectors by using the Radon transform. As the images are not aligned we cannot use the existing measures such as correlation between the regions for computing the similarity. As a result, we need an appropriate similarity measure for computing the similarity between unaligned images.

Generally, it is expected that two intra images (images of the same object) result in high similarity and inter images (images of different objects) result in low similarity. Sometimes, however, intra images can result in low similarity due to relative high non-linear distortion between the images. Therefore, to deal with non-linear distortion, it is better to compute the similarity between images locally. Moreover, for taking global information of images into account for more discriminating information, the global similarity can be computed by using local similarities of regions. The details of computing local and global similarity is discussed below:

Similarity Metric: Computing the similarity between images based on pixel information is difficult as images are not aligned. Hence, a new similarity metric is required that can compute the similarity between images without aligning them and regardless of whether images are full or partial. As discussed in Section 4.3, the proposed Variance Vector is not affected by misalignment of images as it is

rotation invariant. Figure 6.3 shows Variance Vectors of two common regions extracted from the Lena original and rotated images. Both Variance Vectors coincide perfectly with each other. Ideally, the vectors are supposed to be identical and distance between vectors is supposed to be zero, however due to the presence of noise, these can result in a positive distance between vectors. Moreover, a Variance Vector represents the information of a region of an image. Therefore, the distance between Variance Vectors not only can be used for finding the common regions but also can be used for computing the similarity between images. Similarity between images or regions is inversely proportional to the distance between Variance Vectors. The smaller the distance between Variance Vectors; the higher the similarity between images or regions and vice versa.

6.4.3.1 Similarity Computation

The similarity between images can be computed locally and globally as the local information (regions) of images are used. Local similarity is less affected by non-linear distortion present in images. However, local similarity misses the global information of images, therefore it is required to compute global similarity between images. The details of computing local and global similarity is given below:

i) **Local Similarity**: The proposed method matches images by dividing an image into small regions instead of comparing an entire image as a big single region (as in conventional correlation methods). This is because, in partial images, it is difficult to compare entire partial images as they might consist of different parts of the same object. In addition, to lower the effect of non-linear distortion present in images it is better to compute similarity locally as non-linear distortion has less effect on local regions. The distance between Variance Vectors of regions in two images is taken as the local similarity between them.

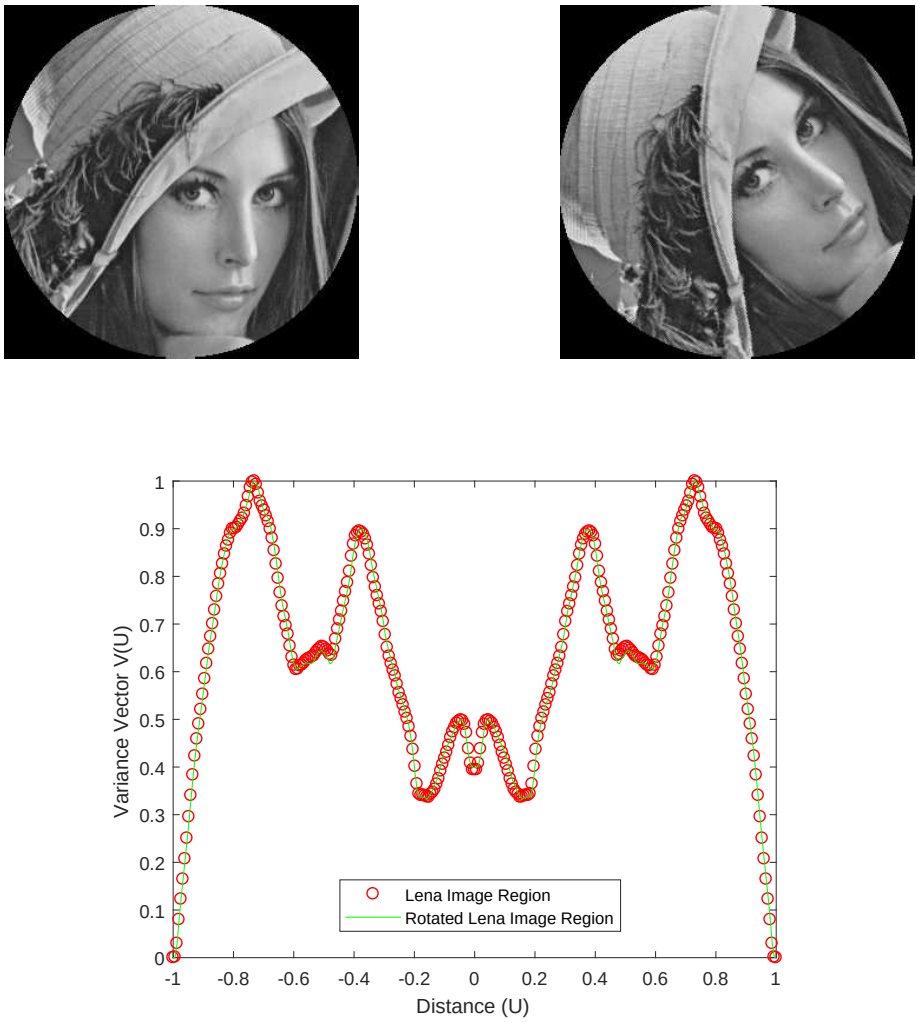


Figure 6.3: The common regions between the Lena original and rotated images and their rotation invariant Variance Vectors

For a higher level of confidence in matching images, more distinguishing information is required that is captured by multiple regions of an image. Multiple regions cover a greater area of an image. This greater area in turn captures more distinguishing information in the image. As discussed in Section 6.4.1, multiple corresponding regions between images can be disjoint or overlapping. As the number of common regions between images increases, the similarity between intra images increases and inter images decreases. This is because, it is possible to have a few common regions between inter images due to inter image similarity. However, it is not possible to have all

regions matching with high similarity between inter images. On the contrary, in intra images it is possible to have more common regions with higher similarity. Hence, for computing similarity between regions, we extend the notion that the distance between Variance Vectors is computed for corresponding multiple regions. The difference between each region pair is considered as a local similarity between them.

ii) **Global Similarity**: Use of the global information of images, increases the distinguishing information of images. Therefore, after computing local similarities of corresponding multiple overlapping regions, a global similarity needs to be computed from the local similarities to indicate the overall similarity between images. There are a number of consolidation techniques and averaging methods that can used to compute global similarity [5] .

Cappeli et al. [5] have discussed different global consolidation techniques that can be used to compute a global similarity value by using the local similarities. A pixel based matching technique discussed in [10] also uses these consolidation techniques, with a few changes to compute the global similarity. Therefore, we can also apply these techniques to compute the global similarity between images.

The simplest consolidation technique is the local similarity sort (LSS) that can be used efficiently on light architectures [5]. This technique sorts the local similarities and selects the best n (where n is a positive integer) number of similarities. The value of n can be decided in two ways (i) select a constant value of n for all images; (ii) select the value of n based on the total number of regions present in the images. In our experiments, for performing analysis we have taken a constant value for n for all images. Let's say, S is the set of

best n local similarities (LS) and is given by:

$$S = \{LS_1, LS_2....., LS_n\} \tag{6.1}$$

Once the best n values of local similarities are computed, the next step is to compute a global similarity value between images. The global similarity value can be computed by averaging the local similarities. There are many averaging methods in the literature such as harmonic mean, geometric mean and arithmetic mean that can be used for averaging [10]. However, the most widely used global consolidation techniques involve averaging the local similarities of the regions by using the arithmetic mean [5].

$$GS = \frac{\sum_{i=1}^{n} LS_i}{n} \tag{6.2}$$

Hence, in the following experimentation, the global similarity between images is computed by averaging the best n local similarities resulting from n common regions of images without any loss of generality. Considering more overlapping regions can increase the amount of distinguishing information which will in turn improve the global similarity value. Based on this global similarity value, the reference image matching the sensed image is identified.

6.5 Performance of the Alignment-free Matching Approach

The absence of pixel correspondence between images makes matching partial images difficult. One of the most widely used applications which encounters the problem of partial images is fingerprint identification. Identification is the process of determining a person or thing with the entity providing only an identifier. The

ever increasing demand for reliable and secure identification of individuals has enabled the use of fingerprints in unprecedented contexts. Moreover, the availability of fingerprints at the time of identification provides great convenience to the users while maintaining security, which in turn prompts the use of fingerprints for identification [3]. In addition, due to increased data handling ability, modern technology support and rising demand for reliable identification, fingerprint identification is also used for civilian and government applications like access control, border security and India's Aadhaar Card project [16]. Inability to identify a person may lead to breach of privacy, lack of access to resources etc.

Fingerprints are the patterns found on the fingers of human beings and are composed of valleys and ridges. The uniqueness of fingerprints is verified by manual inspection and the immutability is proven by morphogenesis of the ridges on the skin [3]. The ever increasing use of fingerprints has given a misconception that there are no issues remaining in an automated fingerprint recognition systems. Despite extensive research in fingerprint identification, there are many issues that still need to be addressed. An example of misidentification is the 2004 Madrid Bomb blast. In 2004, Brandon Mayfield was falsely accused of the train bombing in Madrid that killed 191 people and injured 2000 people [210]. This happened because the fingerprints found at the incident place were wrongly matched to Brandon Mayfield's fingerprint. The fingerprints collected from the crime scene were partial and of low quality[1].

Fingerprints captured in an uncontrolled environment are generally partial and transformed. The matching of such fingerprints by using feature based techniques is difficult. Also, alignment is an obligatory operation in existing identification systems that is exhaustive and time consuming[2]. Maio et al. [211] have shown

[1]Later, when the Spanish government identified the fingerprints and found Brandon Mayfield innocent [210].

[2]It clear that we still need to develop the automated fingerprint identification that identifies the fingerprints efficiently and accurately.

that 20% of poor quality fingerprints in the database are responsible for 80% of the false rejections. Therefore, for partial fingerprint identification, fingerprints must be matched accurately and efficiently under all constraints. Also, it is preferable if the system is accurate, reliable, efficient and effective in a way that works for full, partial and low quality fingerprints. Otherwise the fingerprint based identification will not be the option of choice.

The proposed approach fulfils the conditions required for matching partial fingerprints. Therefore, to verify the accuracy and efficiency of fingerprint identification, we have tested our approach on fingerprint datasets, which are widely used by existing feature and pixel based matching algorithms. The experiments have been performed on four FVC2002 datasets (namely, DB1, DB2, DB3 and DB4) [211] and one FVC2006 dataset [212] which are the most widely used datasets in fingerprint identification [3]. The specifications of these datasets is shown in Figure 6.4. All four datasets of FVC2002DB1 were captured by using different sensors. Each of these datasets consist of fingerprints of 100 individuals with 8 impressions for each individual. Therefore, these datasets consist of 800 fingerprints in total. Volunteers for acquiring these datasets are mainly habituated students and they have been asked to add perturbations deliberately in different impressions. Due to translation, rotation, wetness or dryness of fingers during acquisition, different impressions of the same person are partially overlapping and therefore, these dataset consists of both partial and full fingerprints. Moreover, no effort is made to control the quality of acquired fingerprints. Old sensors have been used for capturing the fingerprints. Also, after acquiring the fingerprints, high quality fingerprints were taken out from the dataset. Therefore, the resulting fingerprints are of low quality. Similarly, the FVC2006DB2 dataset consists of fingerprints of 140 individuals with 12 impressions for each individual. Therefore, it consists of 1680 fingerprints in total. This dataset consists of fingerprints from manual workers and elderly people, having high distortions in fingerprints. To check the robustness of the proposed

Dataset	Size of each Dataset	Notes
FVC2002	100*8	• Volunteers: Students • Quality: No control • Perturbations: Displacements, rotation, wetness and dryness • Consists of partial and low quality images
FVC2006	140*12	• Volunteers: Manual workers and elderly people • Quality: No control • Consists of partial and low quality images

Figure 6.4: Specifications of the FVC2002 and FVC2006 datasets

method against this type of distortion, we have used the FVC2006DB2 dataset. We are using these datasets because the techniques to which we want to compare our method have been evaluated extensively on these datasets and hence provide a benchmark for comparison. Moreover, these datasets are a good combination of all types of fingerprints including full, partial and low quality fingerprints which are required for checking the robustness of the proposed method.

The performance of the proposed method has been measured in terms of accuracy. That represents how many individuals have been identified correctly in entire database. The accuracy is calculated by using the formula given in [10], which is:

$$Accuracy = \frac{CA + CR}{TG + TI} * 100 \tag{6.3}$$

where CA represents the number of genuine comparisons which are correctly accepted, CR is the number of comparisons which are correctly rejected, TG represents the total number of genuine (intra) comparisons, and TI represents the total number of imposter (inter) comparisons. Furthermore, an Equal Error Rate (EER) is also computed for the proposed method, to compare the proposed method with the state of the art. The EER is the rate at a threshold at which the false

acceptance rate and the false rejection rate are equal. The EER is computed as

$$EqualErrorRate = 100 - Accuracy = 100 - \frac{CA + CR}{TG + TI} * 100 \qquad (6.4)$$

For identifying a sensed image, multiple corresponding regions are identified in all intra and inter class images based on the distance between Variance Vectors. The global similarity value is computed by averaging the local similarity values for different regions. The reference image that results in a maximum global similarity value is considered the matching image. If the matching image belongs to the same intra class images then the match is considered as genuine acceptance, otherwise if the matching fingerprint belongs to a inter class image then the match is considered a false acceptance. Accuracy is computed by using the Equation 6.3. Finally, based on the accuracy of the system, the EER is computed by using Equation 6.4. Many factors influence the accuracy of proposed approach. These issues are discusses below:

6.5.1 Effect of the Number of Regions on Accuracy

One of the important factors that affects the accuracy of the proposed approach is the number of regions under consideration. Figures 6.5 and 6.6 show the effect of using multiple overlapping regions on the accuracy of the proposed approach. The accuracy shown in Figures 6.5 and 6.6 are obtained by arbitrarily taking the first impression of each person as a sensed image in the FVC2002DB1 dataset and the sixth impression of each person as a sensed image from the FVC2002DB2 dataset, respectively. To verify the effectiveness of the proposed approach for partial and low quality images, we have taken different impressions as sensed images in both datasets because different impressions consist of different levels of distortion.

The results show that accuracy improves with an increase in the number of regions, because the amount of distinguishing information captured increases with

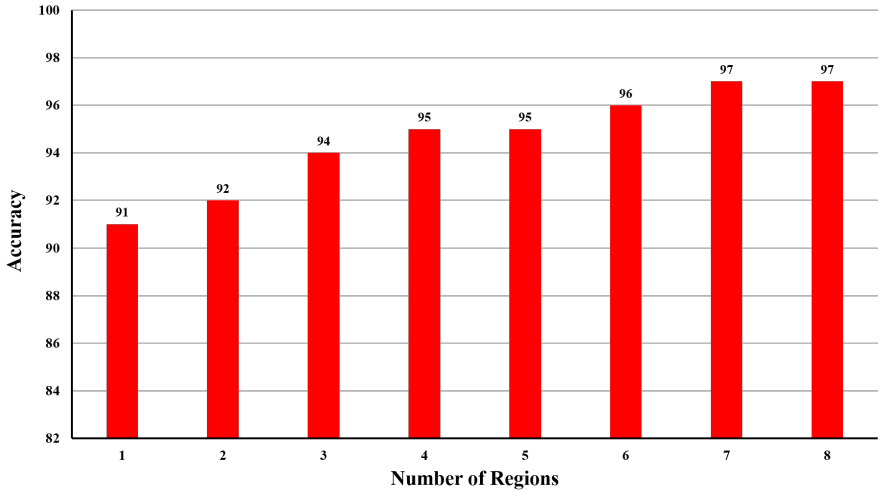


Figure 6.5: The effect of the number of regions on the accuracy achieved by the proposed approach on the FVC2002DB1 dataset. The accuracy is achieved by considering the first impression of each person as a sensed image.

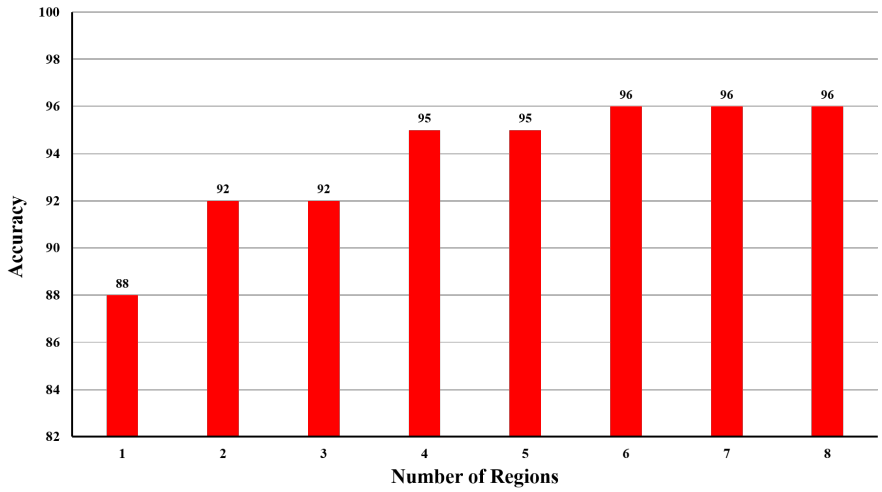


Figure 6.6: The effect of multiple regions on the accuracy achieved by the proposed approach on the FVC2002DB2 dataset. The accuracy is achieved by considering the sixth impression of each person as a sensed image.

number of regions. The accuracy up to eight overlapping regions has been shown. Considering more than 8 regions results in no improvement in accuracy for these two datasets, therefore, 8 regions are considered for the experiment.

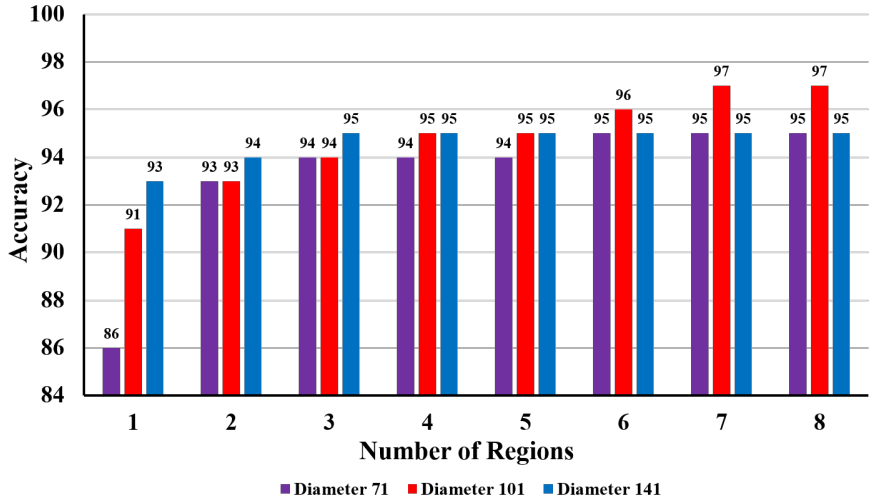


Figure 6.7: The effect of the size of regions on the accuracy of the proposed approach. The accuracy is computed for three different diameters (71, 101 and 141 pixels) of circular regions by considering the first impression of each person as a sensed image for the FVC2002DB1 dataset.

6.5.2 Effect of Size of Region on Accuracy

The next factor that can effect the accuracy of the proposed approach is comprised of the sizes of regions. To study the effect of size of a region, the experiment is performed for regions of 3 different diameters 71, 101 and 141 pixels on the FVC2002DB1, FVC2002DB2, FVC2002DB3, FVC2002DB4 and FVC2006DB2 datasets. By using the method discussed in the Section 6.4.1, the corresponding common matching regions are found between the sensed and reference images. Then the global similarity between images is computed (as discussed in Section 6.4.3). Based on the similarity, the accuracy is computed for all 3 diameters of regions.

Figure 6.7 shows the accuracy achieved by the proposed approach for three different diameters of regions and with increasing number of regions. For this experiment, first impression of each person is taken as sensed image in FVC2002DB1 dataset. Table 6.1 demonstrate the accuracy of proposed approach for three diameters for FVC2002DB1, FVC2002DB2, FVC2002DB3, FVC2002DB4 and FVC2006DB2

datasets. Datasets FVC2002DB2, FVC2002DB3, and FVC2006DB2 are not evaluated for size 141 pixels as many an individual's fingerprints are partial and do not consist of even a single region of size 141 pixels. The results show that the accuracy achieved by the proposed approach increases with an increase in the number of regions (as discussed in Section 6.5.1).

Moreover, the accuracy achieved by a single region of different diameters shows that the larger the region size, the higher the accuracy. This is because, a larger area captures more distinguishing information from fingerprints. Higher distinguishing information between fingerprints results in a higher similarity between them that in turn improves the accuracy [3].

Table 6.1 shows the accuracy achieved by using regions of diameters 71, 101 and 141 pixels. The best accuracy is achieved by a circular region of diameter 101 pixels for different datasets. This is because, when the region size is smaller, it captures less distinguishing information of fingerprints and local similarities of regions will be higher regardless of intra and inter fingerprints. However, if the region size is big then it is quite possible that the sensed image region will not exist in the reference image completely. Also, the larger regions are more prone to non-linear distortions that affect the overall accuracy. Therefore, the best diameter size for circular region among these 3 sizes is 101 pixels, which not only results in maximum accuracy but is also suitable for different datasets.

6.5.3 Performance Evaluation

Different impressions of every person has different quality and distortion [3]. Moreover, some of the impressions are more partial than others. Therefore, to verify the effectiveness of the proposed approach for partial and low quality images, we have taken different impressions of each person as sensed images in all datasets.

[3]This result motivates us to use more regions so that the amount of distinguishing information increases and so does the accuracy.

Datasets	Size of Region (Pixels)	Number of Regions							
		1	2	3	4	5	6	7	8
FVC2002 DB1	71	86	93	94	94	94	95	95	95
	101	91	93	94	95	95	96	97	**97**
	141	93	94	95	95	95	95	95	95
FVC2002 DB2	71	82	88	91	93	93	94	95	96
	101	94	95	97	97	97	97	97	**98**
	141	90	90	96	96	96	97	97	97
FVC2002 DB3	71	68	78	82	82	82	82	83	87
	101	77	88	88	89	89	91	91	**91**
FVC2002 DB4	71	72	84	88	91	92	92	94	95
	101	91	91	93	94	94	94	95	**97**
FVC2006 DB2	71	88.57	95	96.42	96.42	97.14	97.14	97.85	97.85
	101	94.28	96.42	97.14	97.14	97.85	98.57	99.3	**99.3**

Table 6.1: The accuracy achieved by the proposed method on different fingerprint datasets and for three different diameters (71,101,141 pixels) of a circular region by taking one of the impressions as sensed image.

As the previous section shows that circular regions of diameter 101 pixels result in the best accuracy, we have used regions of size 101 pixel diameters for computing the accuracy. The accuracy is computed for multiple overlapping regions and the average results are reported. Figure 6.8 shows the average accuracy obtained by the proposed approach on the FVC2002DB1 datasets. The accuracy is computed for circular regions with 101 pixel diameters with a different impression of each person as a sensed image. We can see that the average accuracy improves with an increase in the number of regions.

Tables 6.2 and 6.3 show the accuracy achieved by taking different impressions as the sensed image for the FVC2002 and FVC2006 datasets, respectively. Moreover, the average accuracy achieved by each dataset is also reported. As can be seen in Tables 6.2 and 6.3, different impressions result in different accuracies because some impressions with respect to the complete fingerprint cover a much larger proportion; and others cover a much smaller proportion. Moreover, at least one impression in each dataset results in 100% accuracy even when the inter

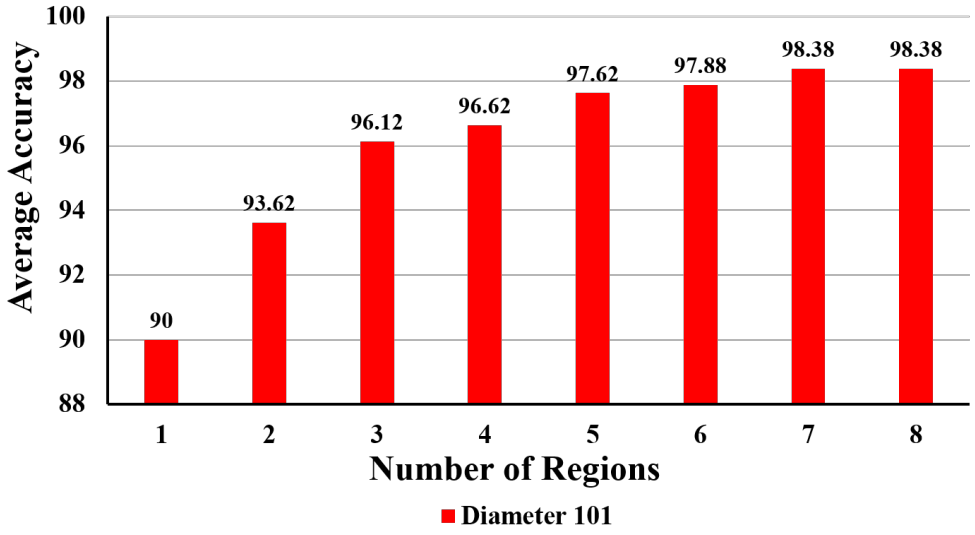


Figure 6.8: The average accuracy with respect to the number of regions on the FVC2002DB1 dataset

similarity exists between fingerprints. The average accuracy achieved by datasets FVC2002DB1, FVC2002DB2, FVC2002DB3, FVC2002DB4 and FVC2006DB2 are 98.38%, 97.75%, 95%, 97.38% and 98.22%, respectively.

Based on the above accuracy, the EER of the proposed method is computed by using Equation 6.4. The EER of the proposed method is compared with the EER of the existing methods for each dataset. Table 6.2 shows the best accuracy obtained by regions of diameter 101 pixels is 98.38% for the FVC2002DB1 dataset, which is far better than the existing techniques. Table 6.4 shows that the EER of the proposed method for FVC2002DB1 results is 1.62%. That is better when compared to the existing techniques. The existing region based approach for partial fingerprints discussed in [213] have achieved the best known EER, 2.66% so far on the FVC2002DB1 dataset, but the proposed method outperforms this method. Moreover, the proposed method performs better than the approaches discussed in [214] and [215] that result in EER of 2.27% and 2% respectively as compared to all other techniques discussed in Table 6.4. The EER of the proposed method on the FVC2002DB2 and FVC2002DB3 datasets are 2.25 and 5, respectively. The method proposed in [215] results in an EER of 2.3 on the FVC2002DB2 dataset;

Dataset	Impression as Sensed Image	Accuracy (DB1)	Accuracy (DB2)	Accuracy (DB3)	Accuracy (DB4)
FVC2002	1	97	98	94	95
	2	100	100	96	97
	3	98	100	99	99
	4	98	97	90	99
	5	98	96	91	93
	6	98	96	93	100
	7	99	97	97	98
	8	99	98	100	98
Average Accuracy		**98.38**	**97.75**	**95.00**	**97.38**

Table 6.2: The accuracy achieved for the FVC2002DB1-DB4 datasets by using the proposed method and by taking different impressions of each person as a sensed image

Dataset	Impression as Sensed Image	Accuracy
FVC2006 DB2	1	95.7
	2	98.6
	3	99.3
	4	99.3
	5	97.9
	6	97.9
	7	99.3
	8	95.7
	9	98.6
	10	100
	11	97.1
	12	99.3
Average Accuracy		**98.22**

Table 6.3: The accuracy achieved for the FVC2006DB2 dataset by using the proposed method and by taking different impressions of each fingerprint as sensed images

Methods (DB1)	**EER (%)**
Qadar, 2007 [111]	7.13
Zhang, 2007 [109]	3.49
Gao, 2011 [82]	3.5
Ahmad, 2011 [216]	9
Das, 2012 [214]	2.27
Wang, 2012 [217]	3.5
Wang, 2014 [215]	2
Zanganeh, 2014 (Conventional) [10]	7.1
Zanganeh, 2014 [10]	2.66
Sandhya, 2015 [218]	4.71
Gowthami, 2015 [219]	3.33
Proposed Method	**1.62**

Table 6.4: Comparison of the EER of the proposed method with the EER of existing methods on the FVC2002DB1 dataset

Methods (DB2)	**EER (%)**
Ahmad, 2011 [216]	6
Wang, 2012 [217]	4
Wang, 2014 [215]	2.3
Sandhya, 2015 [218]	3.44
Gowthami, 2015 [219]	8
Proposed Method	**2.25**

Table 6.5: Comparison of the EER of the proposed approach with the EER of existing methods on the FVC2002DB2 dataset

the method proposed in [219] results in an EER of 5.86 on FVC2002DB3. We can see in Tables 6.5 and 6.6 that there is a small increment in the EER of the proposed method as compared to the existing techniques. However, these datasets consist of highly distorted images. Therefore, a small increment in the EER makes a significant difference. Moreover, the methods proposed in [215] and [219] result in higher EER for other datasets except FVC2002 DB2 and DB3, respectively, whereas the proposed approach results in lower EER for other datasets as well.

In the case of FVC2002DB4, the proposed method improves the EER from 5.5 (using the method proposed in [219]) to 2.62. Not many methods have been evaluated on this dataset, however among the given methods the proposed method results in excellent performance. Similar results in Table 6.8 have been reported

Methods (DB3)	EER (%)
Ahn, 2008 [220]	11.8
Ahmad, 2011 [216]	27
Wang,2012[217]	7.5
Wang, 2014 [215]	6.12
Sandhya,2015 [218]	8.79
Gowthami, 2015 [219]	5.86
Proposed Method	**5**

Table 6.6: Comparison of the EER of the proposed method with the EER of existing methods on the FVC2002DB3 dataset

Methods (DB4)	EER (%)
Ahn, 2008 [220]	11.46
Gowthami, 2015 [219]	5.5
Proposed Method	**2.62**

Table 6.7: Comparison of the EER of the proposed method with the EER of existing methods on the FVC2002DB4 dataset.

for the FVC2006DB2 dataset which consist of distorted fingerprints from manual labourers and elderly people. The proposed method reports 0.78 improvement in the EER. The existing method. proposed in [219] results in a best EER of 2.56 whereas the proposed approach results in an EER of 1.78.

Thus, the proposed approach performs best in comparison with the existing approaches for all datasets even when the datasets consist of partial fingerprints derived from manual labour and elderly people's fingerprints. The proposed method not only achieves the best EER but also simplified the matching process. This is because images are matched without aligning them.

Methods	EER (%)
Abraham, 2011 [221]	11.09
Zanganeh, 2015 (Thesis)	2.56
Proposed Method	**1.78**

Table 6.8: Comparison of the EER of the proposed approach with the EER of existing methods on FVC2006DB2 dataset.

6.5.4 Effect of Noise on Accuracy

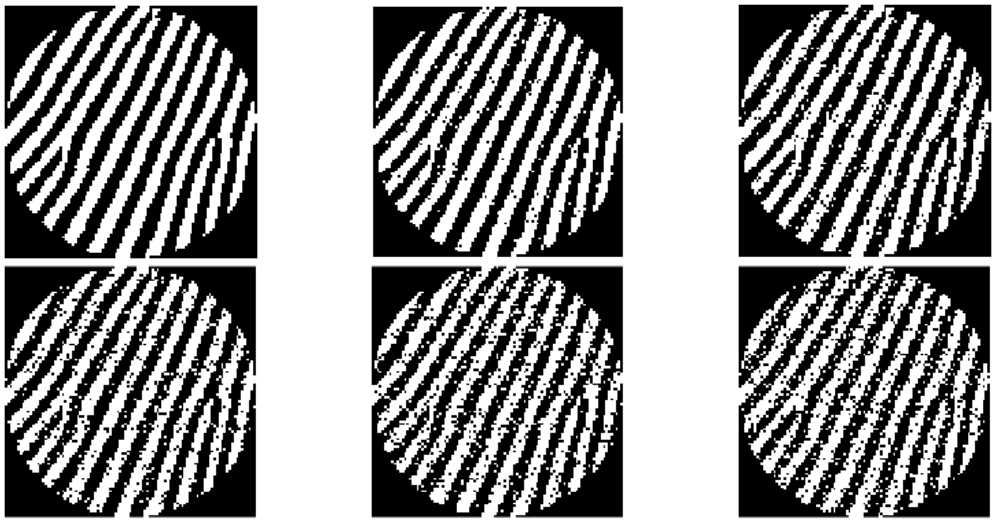


Figure 6.9: Region from a noiseless fingerprint (top left) and its transformed versions obtained by adding salt and pepper noise of varying densities d=0.04,0.08,0.012,0.16,0.20

Noise in fingerprints occurs as random variations of a pixel's brightness. The fingerprints suffer from noise that is generally caused by dirt on the sensor or the degree of dampness of a finger [222]. In noisy fingerprints, the ridges of a finger are not well defined, making identification difficult. Additionally, noise is one of the main causes of intra-class variation in fingerprints. Therefore, estimation of the robustness of the proposed method with respect to additive noise is needed. The experiments have been carried out by adding "salt & pepper" noise to sensed fingerprints (as shown in Figure 6.9) with density d ranging from 0 (noiseless image) to 0.20 in increments of 0.04. The density of noise represents the percentage of pixels corrupted in the images. The noise is added to each fingerprint considered a sensed image. An experiment of fingerprint identification is performed to identify the individuals. The accuracy obtained with different noise levels are shown in Table 6.9.

The results show that accuracy decreases gradually as the amount of additive noise increases. Up to 16% of noise, the accuracy for all datasets is more than 90%.

Datasets	Noise Density (Salt & Pepper)					
	0	**0.04**	**0.08**	**0.12**	**0.16**	**0.20**
FVC2002 DB1	98.38	97.17	96.38	94.88	93.13	90.88
FVC2002 DB2	97.75	93.75	92.38	91.13	90	87.63
FVC2002 DB3	95	94.63	93.5	93.38	90.38	87.88
FVC2002 DB4	97.38	97.38	96.63	95.5	94.13	91.38
FVC2006 DB2	98.22	98.22	98.14	98.14	97.71	96.71

Table 6.9: The accuracy of the proposed approach at different densities of salt and pepper noise on the FVC2002DB1-DB4 and FVC2006DB2 datasets

The accuracy for FVC2002DB1, FVC2002DB4 and FVC2006DB2 is maintained above 90% even at 20% of noise. However, for FVC2002DB2 and FVC2002DB3, accuracy plunges to 87.63% and 87.88% respectively at 20% noise. This is because, the images in these datasets are distorted images. Therefore adding more noise affects the accuracy. The robustness of the method to noise is due to the Radon transform which is robust to additive noise [14]. The proposed method uses the Radon transform to convert image regions into Variance Vectors, thereby reducing the effect of noise on images.

To recapitulate, the results of fingerprint identification show the robustness of the proposed method to additive noise and therefore, the method can be exploited for low quality fingerprints as well.

6.5.5 Effect of the Number of Radon Projections on Accuracy

The results on fingerprint datasets discussed till now are obtained by computing the Radon projections for each angle between 1° to 360° to deal with the rotation difference between fingerprints. The amount of captured information varies with

Datasets	Accuracy (Number of Projections)			
	60	**90**	**180**	**360**
FVC2002DB1	95	98.38	98.38	98.38
FVC2002DB2	91	95.13	96.13	97.75
FVC2002DB3	90	94	94.75	95
FVC2002DB4	90.38	96.88	97.13	97.38
FVC2006DB2	97.68	97.68	98.22	98.22

Table 6.10: The accuracy of the proposed approach with respect to the number of projections on the FVC2002DB1-DB4 and FVC2006DB2 datasets

the number of Radon projections that evidently effect the matching accuracy. We have carried out a sensitivity analysis test for finding how many projections are required (between 1° to 360°) to capture enough distinguishing information of fingerprints so that it can be easily distinguished from inter-class fingerprints. The experiments are performed by computing the Variance Vectors of regions for different numbers of Radon projections. The angles for different projections are uniformly sampled between 1° and 360°.

The regions of 101 pixel diameters are used as they performed best among different sized regions (see details in Section 6.5.2). The experiments are conducted with different numbers of projections for computing the accuracy of fingerprint identification by using the proposed method. We can see in Table 6.10, as the number of projections increases the accuracy of fingerprint identification also improves. However, with only 60 projections, we are able to achieve an accuracy more than 90% for all datasets. The number of projections has decreased by one sixth, but the accuracy has not decreased proportionately. Further increasing the number of projections to 90 can achieve an accuracy close to the accuracy that is achievable by using all 360 projections. Reducing the number of projections required decreases the number of computation steps for each comparison required for fingerprint identification. Therefore, a smaller number of projections not only achieve the highest possible accuracy but also reduce the complexity of the system.

6.5.6 Computational Complexity

One of the important issues in image identification considered is the computational complexity of techniques. Although the accuracy of the proposed method is high and better than the existing techniques, the computational complexity is also desired to be low.

The basis of the proposed method is computation and comparison of the Variance Vectors of sensed and reference image regions. The computation of a Variance Vector depends on the Radon transform. Computing the Radon transform is quite immediate [165]. Moreover, the computing time of a Variance Vector depends on the size of an image as well. We have used regions of images, therefore we compute the time required for computing Variance Vectors for different sized regions. The larger the size, the more time required for computing a vector and vice versa. The time required for computing the Variance Vectors of regions of 71, 101 and 141 pixel diameters is 22.7, 36.4, and 60.4 milliseconds, respectively on a desktop with a 3.5GHz CPU and 64GB RAM.

The time required to compute similarity between two images by one of the existing region based methods discussed in [10] for FVC2002DB1 dataset is approximately 849 seconds on Monash University's High Performance Computing Cluster. However, the proposed approach takes on average only 283 seconds which is approximately one-third of the time required by the existing region based method. This time includes the time required for computing the Variance Vectors of sensed image regions, their comparison with all Variance Vectors of reference image regions and computing similarity between them.

As the proposed method is approximately three times faster than the existing method, it will be very useful in exhaustive applications. One such example is latent fingerprint identification where fingerprints obtained from crime scenes are compared with a dataset consisting of a large number of fingerprints of criminals.

The proposed method will help in finding a suspect accurately and quickly as compared with existing techniques. Moreover, as discussed in Section 6.5.5, reducing the number of projections decreases the computational complexity of the proposed approach.

To recapitulate, the proposed method is not only able to achieve the improved accuracy but also eliminate the need for aligning the fingerprints, which is turn reduces computational complexity.

6.6 Conclusion

This chapter presents a new paradigm for matching the partial images without aligning them. The proposed method is independent of any image features. Therefore, it does not suffer from the problems such as feature extraction errors and insufficient numbers of features. The regions of images capture both micro and macro features of images. That helps in capturing more distinguishing information of images. Also, the regions can be of any size or captured from any part of the fingerprint, making it applicable to both partial and poor quality images. Moreover, the overlapping region approach allows use of as much of the information in images as is available. This approach is tolerant to non-linear distortion because regions represent the local information of the image and non-linear distortion has less effect locally.

Predominantly, alignment is one of the obligatory operations in the existing image identification process which makes the entire process computationally exhaustive. The proposed approach is able to find common matching regions between images without aligning them. That reduces the computational complexity of a system. Moreover, for computing the similarity between the unaligned partial images, a new similarity metric is proposed.

Four FVC2002 datasets and FVC2006DB2 dataset are used as a test bed to evaluate the effectiveness of the proposed image matching method. These datasets are chosen because they consist of low quality and partial images. The proposed approach identify images without aligning them, resulting in the best EER as compared to the existing approaches for all datasets. Moreover, the method's robustness to noise is also tested. Hence, the proposed method is not only accurate and efficient but also robust to noise and applicable to all images regardless of being full or partial.

Chapter 7

Conclusion

Due to the increase in the availability of digital images, the demand for accurate and efficient image registration and matching is increasing in different computer vision applications. Image alignment plays an important role in both image registration and matching applications. Most of the existing alignment methods work well when images are of high quality and have exact pixel-to-pixel correspondence. However, images satisfying these demands cannot always be generated. For example, images captured in an uncontrolled environment (fingerprints captured by smart phone sensors) generally do not have exact pixel-to-pixel correspondence. Moreover, the captured images can be distorted and this results in only a part of the image being useful. The alignment of these images using the existing feature based techniques is difficult. Therefore, to solve this problem, many pixel based techniques have been developed. However, these techniques either are computationally exhaustive or not applicable to images which do not have exact pixel-to-pixel correspondence. These issues of alignment techniques prevent the accurate and efficient registration and matching of images.

This thesis aims to overcome the above limitations by proposing a pixel based technique that can improve the accuracy and efficiency of image registration and matching regardless of pixel-to-pixel correspondence between images. A pixel

based parameter computation method is proposed for accurate and efficient image registration, whereas an alignment-free pixel based matching method is proposed for image matching applications. This has been achieved by converting images into a transformed domain that not only simplifies the misalignment between images but also preserves the information that is present in the pixel domain.

7.1 Contributions of Thesis

In this thesis, four key contributions have been made to the area of image registration and matching. These key contributions are presented below, in the sequence in which they appear in the thesis:

i) A transform based method is proposed to accurately and efficiently compute the transformation parameters between images to align them. The method aligns images having pixel-to-pixel correspondence by using pixel based information only, and does not extract image specific features. In addition, to validate the application independence property of the proposed method, the method's performance is checked on images of varying classes. Moreover, robustness of the proposed method to noise is verified by adding different types of noise to images. The proposed method aligns the images accurately and efficiently as compared to existing techniques.

ii) An alignment-free transform based image matching method is proposed to accurately and efficiently match the images having pixel-to pixel-correspondence. The method can match images without computing all the transformation parameters between them. A new similarity metric is proposed to compute the similarity between unaligned images in transformed domains. As the images are matched without aligning this improves the efficiency of matching process. This method is further improved by reducing the number of Radon

projections. The effectiveness of the proposed method is verified on UMD Logo datasets and results demonstrates outstanding performance of proposed approach as compared to the existing techniques.

iii) To align partial images which do not have exact pixel-to-pixel correspondence, a region based method is proposed that can compute the transformation parameters in such images. As the partial images have only part of an image in common between them, an alignment-free method is proposed to extract the common regions for efficiently computing parameters. Moreover, the pixel-to-pixel correspondence is important in regions for parameter computation, therefore, an appropriate shape of region is proposed. To improve the efficiency of the common region extraction process, an optimization technique is also presented. If required, the extracted common regions can be used for aligning the images which do not have exact pixel-to-pixel correspondence where most of the existing methods fail to perform.

iv) For matching partial images which lack exact pixel-to-pixel correspondence, a region based alignment-free method is proposed. For accurate matching of partial images, a method to use most of the available information by means of multiple regions is provided. Moreover, to minimize the effect of non-linear distortion present in images, the similarity is computed locally. To verify the robustness of local similarity at the global level, a method to compute a global similarity score is presented. The method is evaluated on standard fingerprint datasets and the performance is compared with the existing methods. In addition, the robustness to noise is also verified. The results elucidates the excellent performance of the proposed approach on partial images.

7.2 Future Work

The research's primary aim has been to improve the accuracy and efficiency of image registration and matching for different computer vision applications. Towards that end, pixel based techniques have been proposed that can register and match images regardless of the extent of pixel-to-pixel correspondence between them. This is achieved by converting images into a transformed domain and proposing a similarity metric that can compute similarity between images in the transformed domain. A list of some possible future directions is provided below.

i) The proposed method converts two dimensional pixel based information into one dimensional Variance Vectors for matching images. As images are converted in the vectorized form, which do not contain enough information to restore the images, therefore, these vectors can be used to preserve the privacy of images.

ii) The proposed Variance Vector demands partial images on same scale. A method to compute the Variance Vectors of scaled images is required. This method not only would improve the efficiency of the system but is also applicable to many other computer vision applications such as object recognition in images with multiple objects where objects have different scales.

iii) The proposed alignment-free partial image matching method is tested on fingerprints only. As the method is independent of specific features of images, the method can be extended for other biometrics and computer vision applications.

iv) The proposed parameter computation method is applicable only on images which are transformed by similarity transformation parameters. Indeed, images such as bone, MRI/CT or weld defect images are subjected to these

transformations during acquisition. However, in applications such as medical imaging, images are also subjected to affine transformations like non-uniform scaling and shearing. Thus, extension of the proposed method to affine transformations will be an interesting topic to investigate.

v) A region based matching approach is proposed for matching partial images. For improving the accuracy of the proposed method, the geometry of the extracted common regions can also be considered. The relative positions of regions will improve the accuracy of extracted common regions, which in turn enhance the accuracy of the matching process.

Bibliography

[1] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

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

[3] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*. Springer Science & Business Media, 2009.

[4] S. Chikkerur, A. N. Cartwright, and V. Govindaraju, "K-plet and coupled bfs: a graph based fingerprint representation and matching algorithm," in *Advances in Biometrics*. Springer, 2006, pp. 309–315.

[5] R. Cappelli, M. Ferrara, and D. Maltoni, "Minutia cylinder-code: A new representation and matching technique for fingerprint recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 12, pp. 2128–2141, 2010.

[6] M. Tico and P. Kuosmanen, "Fingerprint matching using an orientation-based minutia descriptor," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 8, pp. 1009–1014, 2003.

[7] S. Leutenegger, M. Chli, and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2548–2555.

[8] F. P. Oliveira and J. M. R. Tavares, "Medical image registration: a review," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 17, no. 2, pp. 73–93, 2014.

[9] R. Szeliski, "Image alignment and stitching: A tutorial," *Foundations and Trends® in Computer Graphics and Vision*, vol. 2, no. 1, pp. 1–104, 2006.

[10] O. Zanganeh, B. Srinivasan, and N. Bhattacharjee, "Partial fingerprint matching through region-based similarity," in *Digital lmage Computing: Techniques and Applications (DlCTA), 2014 International Conference on*. IEEE, 2014, pp. 1–8.

[11] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Radon transform orientation estimation for rotation invariant texture analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 6, pp. 1004–1008, 2005.

[12] A. K. Jain, S. Prabhakar, L. Hong, and S. Pankanti, "Filterbank-based fingerprint matching," *Image Processing, IEEE Transactions on*, vol. 9, no. 5, pp. 846–859, 2000.

[13] A. P. Papliński, "The angular integral of the Radon transform (anirt) as a feature vector in categorization of visual objects," in *Advances in Neural Networks–ISNN 2013*. Springer, 2013, pp. 523–531.

[14] N. Nacereddine, S. Tabbone, and D. Ziou, "Similarity transformation parameters recovery based on Radon transform. application in image registration

and object recognition," *Pattern Recognition*, vol. 48, no. 7, pp. 2227–2240, 2015.

[15] D. Doermann, E. Rivlin, and I. Weiss, "Applying algebraic and differential invariants for logo recognition," *Machine Vision and Applications*, vol. 9, no. 2, pp. 73–86, 1996.

[16] A. Sankaran, M. Vatsa, and R. Singh, "Latent fingerprint matching: A survey." *IEEE Access*, vol. 2, no. 982-1004, p. 1, 2014.

[17] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

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

[19] G. Chen and S. Coulombe, "A new image registration method robust to noise," *Multidimensional Systems and Signal Processing*, vol. 25, no. 3, pp. 601–609, 2014.

[20] J. Zhang, D. Huang, J. Gui, and W. Ye, "2d registration based on contour matching for partial matching images," *Journal of Central South University*, vol. 21, no. 12, pp. 4553–4562, 2014.

[21] D. Ghosh and N. Kaabouch, "A survey on image mosaicing techniques," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 1–11, 2016.

[22] R. D. Eastman, N. S. Netanyahu, and J. Le Moigne, "Survey of image registration methods," in *Image Registration for Remote Sensing*. Cambridge University Press, 2011, pp. 35–78.

[23] D. K. Karna, S. Agarwal, and S. Nikam, "Normalized cross-correlation based fingerprint matching," in *Computer Graphics, Imaging and Visualisation,*

2008. CGIV'08. Fifth International Conference on. IEEE, 2008, pp. 229–232.

[24] A. Bartoli, "Groupwise geometric and photometric direct image registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2098–2108, 2008.

[25] M. V. Wyawahare, P. M. Patil, H. K. Abhyankar *et al.*, "Image registration techniques: an overview," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 2, no. 3, pp. 11–28, 2009.

[26] A. Andreopoulos and J. K. Tsotsos, "50 years of object recognition: Directions forward," *Computer Vision and Image Understanding*, vol. 117, no. 8, pp. 827–891, 2013.

[27] C. Harris and M. Stephens, "A combined corner and edge detector." in *Alvey Vision Conference*, vol. 15. Citeseer, 1988, pp. 147–151.

[28] W. Förstner and E. Gülch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, 1987, pp. 281–305.

[29] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.

[30] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[31] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European Conference on Computer Vision*. Springer, 2006, pp. 404–417.

[32] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *European Conference on Computer Vision*. Springer, 2006, pp. 430–443.

[33] G. Yang, C. V. Stewart, M. Sofka, and C.-L. Tsai, "Automatic robust image registration system: Initialization, estimation, and decision," in *Computer Vision Systems, 2006 ICVS'06. IEEE International Conference on*. IEEE, 2006, pp. 23–23.

[34] W. Burger and M. J. Burge, "Non-rigid image matching," in *Digital Image Processing*. Springer, 2016, pp. 587–608.

[35] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision." in *IJCAI*, vol. 81, no. 1, 1981, pp. 674–679.

[36] A. K. Dewdney, "Analysis of a steepest-descent image-matching algorithm," *Pattern Recognition*, vol. 10, no. 1, pp. 31–39, 1978.

[37] P. Bingham and L. Arrowood, "Projection registration applied to nondestructive testing," *Journal of Electronic Imaging*, vol. 19, no. 3, 2010.

[38] R. Mooser, F. Forsberg, E. Hack, G. Székely, and U. Sennhauser, "Estimation of affine transformations directly from tomographic projections in two and three dimensions," *Machine Vision and Applications*, vol. 24, no. 2, pp. 419–434, 2013.

[39] F. Hjouj and D. W. Kammler, "Identification of reflected, scaled, translated, and rotated objects from their Radon projections," *IEEE Transactions on Image Processing*, vol. 17, no. 3, pp. 301–310, 2008.

[40] L. Li, "Image matching algorithm based on feature-point and DAISY descriptor," *Journal of Multimedia*, vol. 9, no. 6, pp. 829–834, 2014.

[41] D. Le Gall, "Mpeg: A video compression standard for multimedia applications," *Communications of the ACM*, vol. 34, no. 4, pp. 46–58, 1991.

[42] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *European Conference on Computer Vision*. Springer, 1992, pp. 237–252.

[43] M. Irani and P. Anandan, "Video indexing based on mosaic representations," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 905–921, 1998.

[44] M.-C. Lee, W.-G. Chen, C.-l. B. Lin, C. Gu, T. Markoc, S. I. Zabinsky, and R. Szeliski, "A layered video object coding system using sprite and affine motion model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 130–145, 1997.

[45] S. Mann and R. Picard, *Beingundigitalwith digital cameras*. MIT Media Lab Perceptual, 1994.

[46] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 1996, pp. 43–54.

[47] R. Szeliski and H.-Y. Shum, "Creating full view panoramic image mosaics and environment maps," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., 1997, pp. 251–258.

[48] H.-Y. Shum and R. Szeliski, "Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment," *International Journal of Computer Vision*, vol. 36, no. 2, pp. 101–130, 2000.

[49] Y. Jie, Z. Renjie, S. Qifa *et al.*, "Fingerprint minutiae matching algorithm for real time system," *Pattern Recognition*, vol. 39, no. 1, pp. 143–146, 2006.

[50] D. Capel and A. Zisserman, "Automated mosaicing with super-resolution zoom," in *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on.* IEEE, 1998, pp. 885–891.

[51] P. F. McLauchlan and A. Jaenicke, "Image mosaicing using sequential bundle adjustment," *Image and Vision Computing*, vol. 20, no. 9, pp. 751–759, 2002.

[52] M. Brown, D. G. Lowe *et al.*, "Recognising panoramas." in *International Conference on Computer Vision*, vol. 3, 2003, pp. 1218–1227.

[53] M. J. Hannah, "Computer matching of areas in stereo images," Stanford Univ CA Dept Of Computer Science, Tech. Rep., 1974.

[54] ——, "Digital stereo image matching techniques," *International Archives of Photogrammetry and Remote Sensing*, vol. 27, no. B3, pp. 280–293, 1988.

[55] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

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

[57] M. Sester, H. Hild, and D. Fritsch, "Definition of ground-control features for image registration using gis data," *International Archives of Photogrammetry and Remote Sensing*, vol. 32, pp. 538–543, 1998.

[58] A. K. Jain, S. Pankanti, S. Prabhakar, L. Hong, and A. Ross, "Biometrics: a grand challenge," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 2. IEEE, 2004, pp. 935–942.

[59] H. P. Moravec, "The stanford cart and the cmu rover," *Proceedings of the IEEE*, vol. 71, no. 7, pp. 872–884, 1983.

[60] J. T. SHI and C. Tomasi, "C., 1994. good features to track," *IEEE Computer Society*, 1994.

[61] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the Seventh IEEE International Conference on*, vol. 2. Ieee, 1999, pp. 1150–1157.

[62] J. R. Quinlan, *C4. 5: Programs for Machine Learning*. Elsevier, 2014.

[63] I. Zoghlami, O. Faugeras, and R. Deriche, "Using geometric corners to build a 2d mosaic from a set of images," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997, pp. 420–425.

[64] M. A. Oskoei and H. Hu, "A survey on edge detection methods," *University of Essex, UK*, 2010.

[65] J. Kittler, "On the accuracy of the sobel edge detector," *Image and Vision Computing*, vol. 1, no. 1, pp. 37–42, 1983.

[66] T. Tuytelaars and L. Van Gool, "Matching widely separated views based on affine invariant regions," *International Journal of Computer Vision*, vol. 59, no. 1, pp. 61–85, 2004.

[67] T. Kadir, A. Zisserman, and M. Brady, "An affine invariant salient region detector," in *European Conference on Computer Vision*. Springer, 2004, pp. 228–241.

[68] J. L. Crowley and A. C. Parker, "A representation for shape based on peaks and ridges in the difference of low-pass transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 2, pp. 156–170, 1984.

[69] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," *Journal of Applied Statistics*, vol. 21, no. 1-2, pp. 225–270, 1994.

[70] M. Agrawal, K. Konolige, and M. R. Blas, "Censure: Center surround extremas for realtime feature detection and matching," in *European Conference on Computer Vision*. Springer, 2008, pp. 102–115.

[71] Z. Shi and V. Govindaraju, "A chaincode based scheme for fingerprint feature extraction," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 462–468, 2006.

[72] S. Di Zenzo, L. Cinque, and S. Levialdi, "Run-based algorithms for binary image analysis and processing," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 1, pp. 83–89, 1996.

[73] Q. Xiao and H. Raafat, "Fingerprint image postprocessing: a combined statistical and structural approach," *Pattern Recognition*, vol. 24, no. 10, pp. 985–992, 1991.

[74] A. Goshtasby and G. C. Stockman, "Point pattern matching using convex hull edges," *IEEE Transactions on Systems, Man, and Cybernetics*, no. 5, pp. 631–637, 1985.

[75] G. Stockman, S. Kopstein, and S. Benett, "Matching images to models for registration and object detection via clustering," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 3, pp. 229–241, 1982.

[76] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching," Sri International Menlo Park CA Artificial Intellegence Center, Tech. Rep., 1977.

[77] S. Ranade and A. Rosenfeld, "Point pattern matching by relaxation," *Pattern Recognition*, vol. 12, no. 4, pp. 269–275, 1980.

[78] N. K. Ratha, K. Karu, S. Chen, and A. K. Jain, "A real-time matching system for large fingerprint databases," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 18, no. 8, pp. 799–813, 1996.

[79] X. Jiang and W.-Y. Yau, "Fingerprint minutiae matching based on the local and global structures," in *Pattern Recognition. Proceedings. 15th International Conference on*, vol. 2. IEEE, 2000, pp. 1038–1041.

[80] X. Wang, F. Wang, J. Fan, and J. Wang, "Fingerprint classification based on continuous orientation field and singular points," in *Intelligent Computing and Intelligent Systems, 2009. ICIS. IEEE International Conference on*, vol. 4. IEEE, 2009, pp. 189–193.

[81] H.-W. Jung and J.-H. Lee, "Live-scanned fingerprint classification with Markov models modified by GA," *International Journal of Control, Automation and Systems*, vol. 9, no. 5, pp. 933–940, 2011.

[82] Z. Gao, X. You, L. Zhou, and W. Zeng, "A novel matching technique for fingerprint recognition by graphical structures," in *Wavelet Analysis and Pattern Recognition (ICWAPR), 2011 International Conference on.* IEEE, 2011, pp. 77–82.

[83] N. K. Ratha, R. M. Bolle, V. D. Pandit, and V. Vaish, "Robust fingerprint authentication using local structural similarity," in *Applications of Computer Vision, 2000, Fifth IEEE Workshop on.* IEEE, 2000, pp. 29–34.

[84] J. Flusser and T. Suk, "Degraded image analysis: an invariant approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 590–603, 1998.

[85] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, pp. 59–73, 2007.

[86] W. W. Zou and P. C. Yuen, "Very low resolution face recognition problem," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 327–340, 2012.

[87] S. Zhong, K. Xia, X. Yin, and J. Chang, "The representation and simulation for reasoning about action based on colored petri net," in *Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on.* IEEE, 2010, pp. 480–483.

[88] S. Abdelsayed, D. Ionescu, and D. Goodenough, "Matching and registration method for remote sensing images," in *Geoscience and Remote Sensing Symposium, 1995. IGARSS'95.'Quantitative Remote Sensing for Science and Applications', International*, vol. 2. IEEE, 1995, pp. 1029–1031.

[89] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Transactions on Image Processing*, vol. 2, no. 3, pp. 311–326, 1993.

[90] H. Jégou, M. Douze, C. Schmid, and P. Pérez, "Aggregating local descriptors into a compact image representation," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.* IEEE, 2010, pp. 3304–3311.

[91] Y. Ke and R. Sukthankar, "Pca-sift: A more distinctive representation for local image descriptors," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2. IEEE, 2004, pp. II–II.

[92] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, 2005.

[93] S. Se, H.-K. Ng, P. Jasiobedzki, and T.-J. Moyung, "Vision based modeling and localization for planetary exploration rovers," in *Proceedings of International Astronautical Congress*, 2004, pp. 434–440.

[94] B. Fan, F. Wu, and Z. Hu, "Aggregating gradient distributions into intensity orders: A novel local image descriptor," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on.* IEEE, 2011, pp. 2377–2384.

[95] T. Ahonen, J. Matas, C. He, and M. Pietikäinen, "Rotation invariant image description with local binary pattern histogram Fourier features," *Image Analysis*, vol. 5575, pp. 61–70, 2009.

[96] Y. Wang, M. Shi, S. You, and C. Xu, "Dct inspired feature transform for image retrieval and reconstruction," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4406–4420, 2016.

[97] E. Tola, V. Lepetit, and P. Fua, "Daisy: An efficient dense descriptor applied to wide-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 815–830, 2010.

[98] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," *Computer Vision–European Conference on Computer Vision 2010*, pp. 778–792, 2010.

[99] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *Computer Vision (ICCV), 2011 IEEE international conference on*. IEEE, 2011, pp. 2564–2571.

[100] A. Alahi, R. Ortiz, and P. Vandergheynst, "Freak: Fast retina keypoint," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE conference on*. IEEE, 2012, pp. 510–517.

[101] V. Govindu, C. Shekhar, and R. Chellappa, "Using geometric properties for correspondence-less image alignment," in *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, vol. 1. IEEE, 1998, pp. 37–41.

[102] J.-W. Hsieh, H.-Y. M. Liao, K.-C. Fan, M.-T. Ko, and Y.-P. Hung, "Image registration using a new edge-based approach," *Computer Vision and Image Understanding*, vol. 67, no. 2, pp. 112–130, 1997.

[103] C. Shekhar, V. Govindu, and R. Chellappa, "Multisensor image registration by feature consensus," *Pattern Recognition*, vol. 32, no. 1, pp. 39–52, 1999.

[104] A. Ventura, A. Rampini, and R. Schettini, "Image registration by recognition of corresponding structures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 3, pp. 305–314, 1990.

[105] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern recognition*, vol. 26, no. 1, pp. 167–174, 1993.

[106] M. Helm, "Towards automatic rectification of satellite images using feature based matching," in *Geoscience and Remote Sensing Symposium, 1991. IGARSS'91. Remote Sensing: Global Monitoring for Earth Management., International*, vol. 4. IEEE, 1991, pp. 2439–2442.

[107] P. Brivio, A. Della Ventura, A. Rampini, and R. Schettini, "Automatic selection of control-points from shadow structures," *International Journal of Remote Sensing*, vol. 13, no. 10, pp. 1853–1860, 1992.

[108] H. Li, B. 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.

[109] Y. Zhang, X. Yang, Q. Su, and J. Tian, "Fingerprint recognition based on combined features," in *Advances in Biometrics*. Springer, 2007, pp. 281–289.

[110] K. Ito, A. Morita, T. Aoki, T. Higuchi, H. Nakajima, and K. Kobayashi, "A fingerprint recognition algorithm using phase-based image matching for low-quality fingerprints," in *Image Processing, 2005. ICIP 2005*, vol. 2. IEEE, 2005, pp. II–33.

[111] H. A. Qader, A. R. Ramli, and S. Al-Haddad, "Fingerprint recognition using zernike moments." *Int. Arab J. Inf. Technol.*, vol. 4, no. 4, pp. 372–376, 2007.

[112] J. Yang, J. Shin, B. Min, J. Lee, D. Park, and S. Yoon, "Fingerprint matching using global minutiae and invariant moments," in *IEEE Image and Signal Processing, 2008. CISP'08. Congress on*, vol. 4, 2008, pp. 599–602.

[113] L. F. Bastos and J. M. R. Tavares, "Improvement of modal matching image objects in dynamic pedobarography using optimization techniques," in *International Conference on Articulated Motion and Deformable Objects*. Springer, 2004, pp. 39–50.

[114] F. P. Oliveira, J. M. R. Tavares, and T. C. Pataky, "Rapid pedobarographic image registration based on contour curvature and optimization," *Journal of Biomechanics*, vol. 42, no. 15, pp. 2620–2623, 2009.

[115] F. P. Oliveira and J. M. R. Tavares, "Matching contours in images through the use of curvature, distance to centroid and global optimization with order-preserving constraint," pp. 91–110, 2009.

[116] Z. M. Kovács-Vajna, "A fingerprint verification system based on triangular matching and dynamic time warping," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1266–1276, 2000.

[117] Y. Chen, S. C. Dass, and A. K. Jain, "Fingerprint quality indices for predicting authentication performance," in *Audio-and Video-Based Biometric Person Authentication*. Springer, 2005, pp. 160–170.

[118] N. Liu, Y. Yin, and H. Zhang, "A fingerprint matching algorithm based on delaunay triangulation net," in *Computer and Information Technology, 2005. CIT 2005. The Fifth International Conference on*. IEEE, 2005, pp. 591–595.

[119] L. Sha, F. Zhao, and X. Tang, "Improved fingercode for filterbank-based fingerprint matching," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 2. IEEE, 2003, pp. II–895.

[120] A. Lumini and L. Nanni, "Two-class fingerprint matcher," *Pattern Recognition*, vol. 39, no. 4, pp. 714–716, 2006.

[121] S. A. Nene and S. K. Nayar, "A simple algorithm for nearest neighbor search in high dimensions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 9, pp. 989–1003, 1997.

[122] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2. IEEE, 2006, pp. 2161–2168.

[123] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[124] D. Simpson, "Introduction to rousseeuw (1984) least median of squares regression," in *Breakthroughs in Statistics*. Springer, 1997, pp. 433–461.

[125] Z. M. Win and M. M. Sein, "Fingerprint recognition system for low quality images," in *SICE Annual Conference (SICE), 2011 Proceedings of*. IEEE, 2011, pp. 1133–1137.

[126] P. Vijayaprasad, M. N. Sulaiman, N. Mustapha, and R. Rahmat, "Partial fingerprint recognition using support vector machine," *Information Technology Journal*, vol. 9, no. 4, pp. 844–848, 2010.

[127] T.-Y. Jea and V. Govindaraju, "A minutia-based partial fingerprint recognition system," *Pattern Recognition*, vol. 38, no. 10, pp. 1672–1684, 2005.

[128] D. Peralta, M. Galar, I. Triguero, D. Paternain, S. García, E. Barrenechea, J. M. Benítez, H. Bustince, and F. Herrera, "A survey on fingerprint minutiae-based local matching for verification and identification: Taxonomy and experimental evaluation," *Information Sciences*, vol. 315, pp. 67–87, 2015.

[129] J. Lewis, "Fast normalized cross-correlation," in *Vision Interface*, vol. 10, no. 1, 1995, pp. 120–123.

[130] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 57–91, 1996.

[131] A. M. Leroy and P. J. Rousseeuw, "Robust regression and outlier detection," *Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1987*, 1987.

[132] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *IEEE Transactions on Computers*, vol. 100, no. 2, pp. 179–186, 1972.

[133] J. Ashburner, K. J. Friston *et al.*, "Nonlinear spatial normalization using basis functions," *Human Brain Mapping*, vol. 7, no. 4, pp. 254–266, 1999.

[134] J. V. Hajnal, N. Saeed, A. Oatridge, E. J. Williams, I. R. Young, and G. M. Bydder, "Detection of subtle brain changes using sub voxel registration and subtraction of serial mr images." *Journal of Computer Assisted Tomography*, vol. 19, no. 5, pp. 677–691, 1995.

[135] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras, "Variational methods for multimodal image matching," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 329–343, 2002.

[136] J. Orchard, "Globally optimal multimodal rigid registration: An analytic solution using edge information," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 1. IEEE, 2007, pp. I–485.

[137] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[138] B. T. Yeo, M. R. Sabuncu, T. Vercauteren, D. J. Holt, K. Amunts, K. Zilles, P. Golland, and B. Fischl, "Learning task-optimal registration cost functions for localizing cytoarchitecture and function in the cerebral cortex," *IEEE transactions on Medical Imaging*, vol. 29, no. 7, pp. 1424–1441, 2010.

[139] C. Kuglin and D. Hines, "Proceedings of the international conference cybernetics and society," 1975.

[140] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[141] A. Andronache, M. von Siebenthal, G. Székely, and P. Cattin, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Medical Image Analysis*, vol. 12, no. 1, pp. 3–15, 2008.

[142] M. A. Viergever, J. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. Pluim, "A survey of medical image registration–under review," 2016.

[143] P. J. Besl, N. D. McKay *et al.*, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[144] M. Van De Giessen, G. J. Streekstra, S. D. Strackee, M. Maas, K. A. Grimbergen, L. J. Van Vliet, and F. M. Vos, "Constrained registration of the wrist joint," *IEEE Transactions on Medical Imaging*, vol. 28, no. 12, pp. 1861–1869, 2009.

[145] C.-L. Tsai, C.-Y. Li, G. Yang, and K.-S. Lin, "The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence," *IEEE Transactions on Medical Imaging*, vol. 29, no. 3, pp. 636–649, 2010.

[146] M.-s. Pan, J.-t. Tang, Q.-s. Rong, and F. Zhang, "Medical image registration using modified iterative closest points," *International Journal for Numerical Methods in Biomedical Engineering*, vol. 27, no. 8, pp. 1150–1166, 2011.

[147] E. Bayro-Corrochano and J. Rivera-Rovelo, "The use of geometric algebra for 3d modeling and registration of medical data," *Journal of Mathematical Imaging and Vision*, vol. 34, no. 1, pp. 48–60, 2009.

[148] G. K. Matsopoulos, P. A. Asvestas, N. A. Mouravliansky, and K. K. Delibasis, "Multimodal registration of retinal images using self organizing maps," *IEEE Transactions on Medical Imaging*, vol. 23, no. 12, pp. 1557–1563, 2004.

[149] D. Maksimov, J. Hesser, C. Brockmann, S. Jochum, T. Dietz, A. Schnitzer, C. Duber, S. O. Schoenberg, and S. Diehl, "Graph-matching based CTA," *IEEE Transactions on Medical Imaging*, vol. 28, no. 12, pp. 1940–1954, 2009.

[150] S. Cho, Y. Chung, and J. Lee, "Automatic image mosaic system using image feature detection and taylor series." in *DICTA*. Citeseer, 2003, pp. 549–560.

[151] S. Chikkerur, V. Govindaraju, S. Pankanti, R. Bolle, and N. Ratha, "Novel approaches for minutiae verification in fingerprint images," in *Application of Computer Vision, 2005. WACV/MOTIONS'05. Seventh IEEE Workshops on*, vol. 1. IEEE, 2005, pp. 111–116.

[152] V. A. Sujan and M. P. Mulqueen, "Fingerprint identification using space invariant transforms," *Pattern Recognition Letters*, vol. 23, no. 5, pp. 609–619, 2002.

[153] L. Lucchese, G. Doretto, and G. M. Cortelazzo, "A frequency domain technique for range data registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, pp. 1468–1484, 2002.

[154] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, 2002.

[155] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 5, pp. 700–703, 1987.

[156] Q.-s. Chen, M. Defrise, and F. Deconinck, "Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1156–1168, 1994.

[157] 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, 1996.

[158] G. Wolberg and S. Zokai, "Robust image registration using log-polar transform," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 1. IEEE, 2000, pp. 493–496.

[159] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.

[160] J. Orchard, "Efficient least squares multimodal registration with a globally exhaustive alignment search," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2526–2534, 2007.

[161] F. P. Oliveira, T. C. Pataky, and J. M. R. Tavares, "Registration of pedobarographic image data in the frequency domain," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 13, no. 6, pp. 731–740, 2010.

[162] R. Xu and Y.-W. Chen, "Wavelet-based multiresolution medical image registration strategy combining mutual information with spatial information," *International Journal of Innovative Computing, Information and Control*, vol. 3, no. 2, pp. 285–296, 2007.

[163] N. Nacereddine, S. Tabbone, D. Ziou, and L. Hamami, "Shape-based image retrieval using a new descriptor based on the Radon and Wavelet transforms," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 1997–2000.

[164] Y. Wan and N. Wei, "A fast algorithm for recognizing translated, rotated, reflected, and scaled objects from only their projections," *IEEE Signal Processing Letters*, vol. 17, no. 1, pp. 71–74, 2010.

[165] N. Nacereddine, S. Tabbone, and D. Ziou, "Object recognition using Radon transform-based rst parameter estimation," in *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer, 2012, pp. 515–526.

[166] G. Yu, W. Cao, and Z. Li, "Rotation and scale invariant for texture analysis based on Radon transform and Wavelet transform," in *Pervasive Computing and Applications, 2008. ICPCA 2008. Third International Conference on*, vol. 2. IEEE, 2008, pp. 704–708.

[167] M. R. Hejazi and Y.-S. Ho, "A hierarchical approach to rotation-invariant texture feature extraction based on Radon transform parameters," in *Image Processing, 2006 IEEE International Conference on*. IEEE, 2006, pp. 1469–1472.

[168] P. Cui, J. Li, Q. Pan, and H. Zhang, "Rotation and scaling invariant texture classification based on Radon transform and multiscale analysis," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 408–413, 2006.

[169] S. Xiao and Y. Wu, "Rotation-invariant texture analysis using Radon and Fourier transforms," *Chinese Optics Letters*, vol. 5, no. 9, pp. 513–515, 2007.

[170] M. Miciak, "Character recognition using Radon transformation and principal component analysis in postal applications," in *Computer Science and Information Technology, 2008. IMCSIT 2008. International Multiconference on.* IEEE, 2008, pp. 495–500.

[171] T. Sandhan, H. J. Chang, and J. Y. Choi, "Abstracted Radon profiles for fingerprint recognition," in *Image Processing (ICIP), 2013 20th IEEE International Conference on.* IEEE, 2013, pp. 4156–4160.

[172] P. Viola and W. M. Wells III, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[173] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. Van Saarloos, "Registration of stereo and temporal images of the retina," *IEEE Transactions on Medical Imaging*, vol. 18, no. 5, pp. 404–418, 1999.

[174] C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[175] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill *et al.*, "A comparison of similarity measures for use in 2-d-3-d medical image registration," *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, pp. 586–595, 1998.

[176] A. Jain, A. Ross, and S. Prabhakar, "Fingerprint matching using minutiae and texture features," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 3. IEEE, 2001, pp. 282–285.

[177] A. Ross, A. Jain, and J. Reisman, "A hybrid fingerprint matcher," *Pattern Recognition*, vol. 36, no. 7, pp. 1661–1673, 2003.

[178] C. Beleznai, H. Ramoser, B. Wachmann, J. Birchbauer, H. Bischof, and W. Kropatsch, "Memory efficient fingerprint verification," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 2. IEEE, 2001, pp. 463–466.

[179] K. Nandakumar and A. K. Jain, "Local correlation-based fingerprint matching." in *Indian Conference on Computer Vision, Graphics and Image Processing*, 2004, pp. 503–508.

[180] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.

[181] P. M. Hambalık, "Fingerprint recognition system using artificial neural network as feature extractor: design and performance evaluation," *Tatra Mt. Math. Publ*, vol. 67, pp. 117–134, 2016.

[182] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, pp. 4–19, 2014.

[183] A. B. Abche, F. Yaacoub, A. Maalouf, and E. Karam, "Image registration based on neural network and Fourier transform," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE.* IEEE, 2006, pp. 4803–4806.

[184] S. Wang, S. Lei, and F. Chang, "Image registration based on neural network," in *Information Technology and Applications in Biomedicine, 2008. ITAB 2008. International Conference on.* IEEE, 2008, pp. 74–77.

[185] S. Bianco, M. Buzzelli, D. Mazzini, and R. Schettini, "Deep learning for logo recognition," *Neurocomputing*, vol. 245, pp. 23–30, 2017.

[186] S. Romberg, L. G. Pueyo, R. Lienhart, and R. Van Zwol, "Scalable logo recognition in real-world images," in *Proceedings of the 1st ACM International Conference on Multimedia Retrieval*. ACM, 2011, pp. 25–32.

[187] S. Romberg and R. Lienhart, "Bundle min-hashing for logo recognition," in *Proceedings of the 3rd ACM Conference on International Conference on Multimedia Retrieval*. ACM, 2013, pp. 113–120.

[188] J. Teng, S. Wang, J. Zhang, and X. Wang, "Neuro-fuzzy logic based fusion algorithm of medical images," in *Image and Signal Processing (CISP), 2010 3rd International Congress on*, vol. 4. IEEE, 2010, pp. 1552–1556.

[189] J. Feng, "Combining minutiae descriptors for fingerprint matching," *Pattern Recognition*, vol. 41, no. 1, pp. 342–352, 2008.

[190] H.-W. Jung and J.-H. Lee, "Noisy and incomplete fingerprint classification using local ridge distribution models," *Pattern Recognition*, vol. 48, no. 2, pp. 473–484, 2015.

[191] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[192] J.-W. Hsieh, H.-Y. M. Liao, K.-C. Fam, and M.-T. Ko, "A fast algorithm for image registration without predetermining correspondences," in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, vol. 1. IEEE, 1996, pp. 765–769.

[193] D. Lavine, B. A. Lambird, and L. N. Kanai, "Recognition of spatial point patterns," *Pattern Recognition*, vol. 16, no. 3, pp. 289–295, 1983.

[194] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the individuality of fingerprints," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 8, pp. 1010–1025, 2002.

[195] S. R. Deans, *The Radon transform and some of its applications.* Courier Corporation, 2007.

[196] S. M. Ross, *Introductory statistics.* Academic Press, 2017.

[197] T. T. Soong, *Fundamentals of probability and statistics for engineers.* John Wiley & Sons, 2004.

[198] M. Hasegawa and S. Tabbone, "Amplitude-only log Radon transform for geometric invariant shape descriptor," *Pattern Recognition*, vol. 47, no. 2, pp. 643–658, 2014.

[199] T. V. Hoang and S. Tabbone, "Invariant pattern recognition using the rfm descriptor," *Pattern Recognition*, vol. 45, no. 1, pp. 271–284, 2012.

[200] A. A. Goshtasby, *2-D and 3-D image registration: for medical, remote sensing, and industrial applications.* John Wiley & Sons, 2005.

[201] K. Komal, D. Albrecht, N. Bhattacharjee, and B. Srinivasan, "A region-based alignment-free partial fingerprint matching," in *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media.* ACM, 2016, pp. 63–70.

[202] 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.

[203] Y. W. Chen and Y. Q. Chen, "Invariant description and retrieval of planar shapes using Radon composite features," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4762–4771, 2008.

[204] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825–848, 2002.

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

[206] K. Komal, N. Bhattacharjee, D. Albrecht, and B. Srinivasan, "Transformational approach for alignment-free image matching applications," in *Proceedings of the 15th International Conference on Advances in Mobile Computing & Multimedia*. ACM, 2017, pp. 49–57.

[207] A. P. Psyllos, C.-N. E. Anagnostopoulos, and E. Kayafas, "Vehicle logo recognition using a sift-based enhanced matching scheme," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 322–328, 2010.

[208] N. Hagbi, O. Bergig, J. El-Sana, and M. Billinghurst, "Shape recognition and pose estimation for mobile augmented reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 10, pp. 1369–1379, 2011.

[209] T.-Y. Jea, V. S. Chavan, V. Govindaraju, and J. K. Schneider, "Security and matching of partial fingerprint recognition systems," in *Proceedings of SPIE*, vol. 5404, 2004, pp. 39–50.

[210] U. S. D. of Justice. Office of the Inspector General, *A Review of the FBI's Handling of the Brandon Mayfield Case*. US Department of Justice, Office of the Inspector General, Oversight and Review Division, 2006.

[211] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "Fvc2002: Second fingerprint verification competition," in *Pattern recognition, 2002. Proceedings. 16th International Conference on*, vol. 3. IEEE, 2002, pp. 811–814.

[212] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni, "Fingerprint verification competition 2006," *Biometric Technology Today*, vol. 15, no. 7-8, pp. 7–9, 2007.

[213] O. Zanganeh, N. Bhattacharjee, and B. Srinivasan, "Partial fingerprint alignment and matching through region-based approach," in *Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia*. ACM, 2015, pp. 275–284.

[214] P. Das, K. Karthik, and B. C. Garai, "A robust alignment-free fingerprint hashing algorithm based on minimum distance graphs," *Pattern Recognition*, vol. 45, no. 9, pp. 3373–3388, 2012.

[215] S. Wang and J. Hu, "Design of alignment-free cancelable fingerprint templates via curtailed circular convolution," *Pattern Recognition*, vol. 47, no. 3, pp. 1321–1329, 2014.

[216] T. Ahmad, J. Hu, and S. Wang, "Pair-polar coordinate-based cancelable fingerprint templates," *Pattern Recognition*, vol. 44, no. 10, pp. 2555–2564, 2011.

[217] S. Wang and J. Hu, "Alignment-free cancelable fingerprint template design: A densely infinite-to-one mapping (ditom) approach," *Pattern Recognition*, vol. 45, no. 12, pp. 4129–4137, 2012.

[218] M. Sandhya and M. V. Prasad, "k-nearest neighborhood structure (k-nns) based alignment-free method for fingerprint template protection," in *Biometrics (ICB), 2015 International Conference on*. IEEE, 2015, pp. 386–393.

[219] A. Gowthami and H. Mamatha, "Fingerprint recognition using zone based linear binary patterns," *Procedia Computer Science*, vol. 58, pp. 552–557, 2015.

[220] D. Ahn, S. G. Kong, Y.-S. Chung, and K. Y. Moon, "Matching with secure fingerprint templates using non-invertible transform," in *Image and Signal Processing, 2008. CISP'08. Congress on*, vol. 2. IEEE, 2008, pp. 29–33.

[221] J. Abraham, P. Kwan, and J. Gao, "Fingerprint matching using a hybrid shape and orientation descriptor," in *State of the Art in Biometrics*. InTechOpen, 2011.

[222] G. Nagendra, V. Sridhar, M. Prameelamma, and M. Zubair, "Fingerprint image enhancement using filtering techniques," *International Journal of Computer Science Engineering*, vol. 1, no. 1, pp. 61–78, 2012.