

CAMERA BASED PALMPRINT RECOGNITION

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (by Research)
in
Computer Science

by

Chhaya Methani

200402010

chhaya@research.iiit.ac.in



International Institute of Information Technology
Hyderabad, India
August 2010

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled “Camera Based Palmprint Recognition” by Chhaya Methani, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date

Advisor: Dr. Anoop. M. Namboodiri

Copyright © Chhaya Methani, 2010
All Rights Reserved

To my parents

Acknowledgements

I would like to thank my advisor Dr. Anoop Namboodiri for the incessant support provided by him through the entire duration of my thesis. The freedom of thought and work style provided by him has helped me achieve a sense of responsibility and being in control of my work. This, I believe, is very important for growth as a researcher. His ability to get quality work done easily under pressure of deadline is truly inspirational. I have been able to look up to him for guidance and open discussions on academic matters outside the scope of my thesis. I thank him for all the above.

This acknowledgement would not be complete without thanking all the people without whom this thesis would not have been possible. First off, I want to mention Himanshu Arora and Sachin Gupta, fellow labmates under Dr. Anoop, with whom I had the pleasure of working shortly on different aspects of my work. I am grateful to both of you for your time and interest in the various little nuances of the project. I want to thank Sheetal for helping out with all the last minute paper revisions and figure alterations in the wee hours of morning, when there was no other soul to be seen in the lab, at the time of submitting PReMI. I am grateful to have amazing seniors whose precious advice and experience I could always draw upon. I thank Pooja and Pradhee Tandon, Gopal Datt Joshi who were always generous with valuable inputs. I will always fondly remember my labmates and friends who have been great fun to be with and made Masters a memorable experience. I thank Maneesh, Suhail, Chandrika, Mihir, Wasif, Pavan D., Chetan, Jyotish, Jinnesh, Venkata Rasagna, Sreekanth, Rakesh, Dileep, and Naveen. I thank Pratyush Bhatt for helping out with various technical aspects many a times. Warm thanks to all those people who took out time to provide palm data for my project. Its a shame that it is not possible for me to include all your names in this short space.

I am eternally grateful to my parents for being patient throughout the course of my Masters degree, even at times when they could not fully comprehend the various delays. My pillars of strength, who would patiently listen to and address my doubts and were always there to boost my spirits on tough days - Varun, Prachi, Sheetal and my younger brother Jai. Thank you guys! I can not imagine what I would have done without you all. I also thank close friends Ranjitha, Vishwas Rao, Pinky, Samba, Dhruvan, Nishant and Aniket for their support at various points in time. Last, but not the least, I want to thank Wasif for helping out with the printing of this thesis.

Abstract

There was a time when biometrics was looked upon as the science of the future. It has featured prominently in various science fiction movies as an advanced security measure used to safeguard important documents, buildings etc. With the help of fast paced technological innovation, today, this is not far from reality. Biometrics is increasingly being used for secure authentication of individuals and making its presence felt in our lives. It uses an individual's physical or behavioral traits to identify them. The decision of which biometric is to be used for a particular application is a complex function of the people's security needs, ease of use and size of the enterprise. People can now see biometrics based security checks at airports that use iris and hand geometry based authentication and at ATMs using fingerprint and hand veins for authentication.

The stage is now set for the use of biometric recognition in commercially viable civilian applications. Mobile devices that can connect to computational servers and laptops can benefit from this technology in making our systems more secure. Most of these devices now come with attached commodity cameras which can be used for biometric image capture. An easy-to-capture biometric modality that could work well with a commodity camera is palmprint. It has coarse lines which can be easily detected using a low resolution camera and it is easy to present due to the free mobility of our palm. On most surveys, hand as a biometric modality rates high on user acceptance. For these reasons, palmprint would be an ideal choice for recognition using commodity cameras.

Consumer devices, however, employ image capture in an open environment as opposed to the controlled environment preferred for biometric data capture by state-of-art designs. For palmprint image capture, a semi-closed environment is created using a box-like setup having an illumination source on top, resulting in clean images with pre-fixed pose and illumination settings. In spite of the resulting high accuracy, it is not practically feasible to provide a semi closed setup with a laptop or a mobile device where people are using their systems everyday and want fast access on a regular basis. The unrestricted imaging associated with mobile cameras results in huge intra class variations of palm. The performance of existing techniques for palmprint authentication fall considerably, when the camera is not aligned with the surface of the palm. The problems arise primarily due to variations in appearance introduced due to varying pose, but is compounded by specularities of the skin and blur due to motion and focus. Hence, we need novel recognition methods for images captured in an unconstrained environment with the hand presented to the system in an unsupervised manner.

In this thesis, we propose the design of a biometric system used for unconstrained and unsupervised camera based palmprint recognition system. We present a new pose transformation algorithm that can identify individuals even after seeing the hand being presented in a pose different than that stored in the database. The method can robustly estimate and correct variations in pose, and compute a similarity measure between the corrected test image and a reference image. The method is able to correct for pose variation even in degraded images having variable illumination.

However, another challenge surfaces during matching of the palm wherein bad line visibility worsens recognition performance. Hence, we propose another algorithm that separates out the original palm lines from the false line like impressions created due to illumination, contrast variations and loose skin. Even minor changes in pose of the palm can induce significant changes in the visibility of the lines. We turn this property to our advantage by capturing a short video, where the

natural palm motion induces minor pose variations, providing additional texture information. This is important for improved matching of images. We propose a method to register multiple frames of the video without requiring correspondence, while being efficient.

Since, this is the first attempt at creating an unconstrained palmprint recognition system, we created two in-house databases to model the pose and illumination variations related to the palm image capture process used by us. The first database contains images of 100 users, having 5 images each having variable poses. The second database captures 6 videos each for 100 subjects captured using a regular web camera. Both the datasets have been captured under natural illumination conditions. Experimental results on the first dataset using the pose correction algorithm shows a reduction in Equal Error Rate from 22.4% to 8.7%. Through an independent experiment performed on the video palm database, we observed that the use of multiple frames reduces the error rate from 12.75% to 4.7%. We also propose a method for detection of poor quality samples due to specularities and motion blur, which further reduces the EER to 1.8%.

Contents

1	Introduction	1
1.1	Introduction	1
1.2	Biometrics - An Overview	3
1.2.1	Classes of biometric systems	4
1.3	Biometric Recognition Process	7
1.3.1	Biometrics: Modes of Operation	7
1.4	Performance Metrics	7
1.4.1	The Basic Model	8
1.4.2	ROC Space	10
1.4.3	Equal Error Rate (EER)	13
1.4.4	Others	14
1.5	An overview of Biometric Technologies	15
1.5.1	Fingerprints	15
1.5.2	Face	16
1.5.3	Iris	17
1.5.4	Retina	17
1.5.5	Gait Recognition	18
1.5.6	Keystroke Dynamics	18
1.5.7	Handwriting Based Recognition	18
1.5.8	Voice Recognition	19
1.5.9	Genetic Fingerprinting	20
1.5.10	Hand Based Recognition	20
1.6	Choosing a Biometric Modality	21
1.7	Introduction to Thesis	22
1.8	Towards Camera Based Biometric Data Capture	22
1.9	Problem Statement	25
1.9.1	Palmprint as a biometric modality	26
1.10	Contributions of the thesis	26
1.11	Thesis Overview	27
2	Literature Survey	28
2.1	Palmprint Recognition	28
2.1.1	Image Capture	29
2.1.2	Image Preprocessing	29
2.1.3	Feature Extraction and Matching	34

2.2	Towards Unconstrained Camera Based Palmprint Imaging	36
2.3	Unconstrained Palm Imaging: Challenges and existing solutions	37
2.3.1	Background Clutter	38
2.3.2	Feature Extraction	39
2.3.3	Scale Variations	39
2.3.4	Pose Variation in images	39
2.3.5	Illumination Variations and Noisy Artifacts	40
2.4	Unconstrained Palmprint Imaging: Bottlenecks	40
2.4.1	Solving Pose Variation for the Palmprint	44
2.4.2	Illumination variations	45
2.5	Summary	46
3	Pose Invariant Palmprint Recognition	47
3.1	Introduction	47
3.2	Previous Work	48
3.3	Unconstrained Palmprint Imaging	51
3.3.1	Finding projective invariant features in palm	51
3.3.2	Pose correction	52
3.4	Image Alignment	52
3.5	Proposed solution	54
3.5.1	Preprocessing	54
3.5.2	Extraction of ROI	56
3.5.3	Image Alignment	57
3.5.4	Feature Extraction	57
3.5.5	Matching	57
3.6	Experimental Results and Discussion	57
3.6.1	Synthetic Dataset	58
3.6.2	Real Dataset	58
3.7	Conclusion and Future Work	59
4	Video Based Palmprint Recognition	62
4.1	Introduction	62
4.2	Combination of Multiple Frames	64
4.3	Capturing Palm Video	65
4.4	Palm Extraction	65
4.5	Registration and Frame Combination	65
4.6	Dataset, Experimental Results and Discussions	68
4.7	Conclusion	72
5	Conclusions and Future Work	74
Appendix		76
5.1	Gabor Filter	76
5.1.1	1D Gabor filter	76
5.1.2	2D Gabor filter	76

5.2	RANSAC Based Homography	78
5.2.1	Homography	78
5.2.2	RANSAC	79

Related Publications	81
-----------------------------	-----------

List of Figures

1.1	Various applications of unconstrained palmprint imaging	2
1.2	Ways of Authentication	6
1.3	Modes of Operation	6
1.4	A Biometric Recognition System.	8
1.5	ROC space and plots of the four prediction examples	11
1.6	ROC curve for a univariate two class probability distribution. Each point on a ROC curve corresponds to a particular confusion matrix at a particular threshold value t_0	14
1.7	Probability distribution of the two class data points in Biometrics, namely the genuine and the imposter matching scores.	15
1.8	Facial Recognition	16
1.9	Fingerprint Based Recognition	16
1.10	Iris and Retina Recognition.	16
1.11	Keystroke Based Recognition	19
1.12	Signature	19
1.13	Gait	19
1.14	DNA Profiling	19
1.15	Speaker Recognition	19
1.16	Hand Based Biometric Recognition	21
1.17	Effect of EER on user acceptance of various Biometric System.	23
2.1	The main steps of preprocessing and palm extraction. a) Original Image, b) Binarized image, c) Boundary Tracking, d) Fixing the co-ordinate system, e) extracting the central region as ROI from the image, and f) resulting extracted palm	30
2.2	The main steps of the modified preprocessing method. (a)binary image, (b)Rotation basing on outer boundary direction and obtain the edge image. (c)truncate the boundary image with 8 intersections (d) trace finger gaps (e) re-rotation of the original image based on the coordinate based on gaps between fingers (f) extract a 128 128 sub-image.	31
2.3	The three approaches for fixing co-ordinate axes for Palm Extraction.	32
2.4	Figure a) and b) show the difference between various palm sizes. Figure c) shows the elliptical palm rings extracted from the hand.	32
2.5	Different image capturing setups. a) Constrained image capture, b) Touchless Imaging, c) Proposed unconstrained touchless capture	36
2.6	Different palm images obtained by using different setups. a) Inked palm, b) scanned palm, c) palm image obtained by unconstrained imaging.	36
2.7	Challenging cases of unconstrained camera imaging	38

2.8	Segmenting hand image from a cluttered background.	38
3.1	Variations in hand pose due to unconstrained imaging.	48
3.2	Extraction of palm region based on finger valleys.	50
3.3	The process being followed to get interest points from the image. a) Palm Image, b) Edge map with a lower threshold, c) Edge map with a higher threshold, d) Edge map after enhancing image in c, e) Sampled interest points marked on the palm image.	53
3.4	Result of aligning Image a to Image b, using Homography, has been shown in Figure c.	54
3.5	a) Flowchart of state of the art palmprint recognition algorithm. b) Proposed algorithm.	56
3.6	A set of images from the Dataset. First row shows 4 images each from 2 users in Synthetic Dataset. 4 left hand images each from 2 users in Real Dataset are shown in the second row and right hand is shown in third row.	58
3.7	ROC curves for the Synthetic and Real Datasets in <i>linear</i> and <i>logarithmic</i> scales.	60
3.8	Images from our dataset with blur, specular reflection, and skin deformations.	61
4.1	Variations in Palm Lines with view variation. The highlighted regions shown separately show that the lines look phased out in different views. c) and e) show clear views.	63
4.2	The process of Registration: a) Line map of the first image, b) Line map of the second image, c) Euclidean Distance Transform(DT) of the first image, d) Gradient Transform of DT, e) overlapped image	63
4.3	gabor response on successive addition of frames.a) gabor response of the Base Image(BI), b)On adding 2 images to BI, c) On adding 6 images to BI, and d) On adding 10 images to BI	64
4.4	Few frames from a video in our dataset showing variations due to illumination and problems in image focusing	64
4.5	ROC curves on <i>logarithmic</i> scale for a) <i>Parameter Set₁</i> , b) <i>Parameter Set₂</i> , c) <i>Parameter Set₃</i> , d) <i>Parameter Set₄</i>	68
4.6	ROC curves on a) <i>linear</i> and b) <i>logarithmic</i> scales, and the result of removing poor captures.	73
5.1	a) an input signal, b) output of Gabor filter(cosine carrier) and c) output of Gabor filter(sine carrier)	77
5.2	The real and imaginary parts of a Gabor filter.	77
5.3	Planar points as viewed by two different cameras	79

List of Tables

- 1.1 Confusion Matrix 9
- 1.2 Confusion matrices corresponding to different classification or prediction systems. 12

- 3.1 Equal Error Rate(EER) on two datasets with the fixed pose and proposed approach. 59

- 4.1 Equal Error Rate(EER) observed using different gabor filter parameters. 69
- 4.2 EER and FTA (in %) with varying number of frames and different combination strategies. The results have been computed for 30 users. 69
- 4.3 EER and FTA (in %) with varying number of frames and different quality thresholds. 71

Chapter 1

Introduction

1.1 Introduction

In recent times, automatic biometric authentication has emerged as the backbone of the new-age solutions to our society's ever increasing demand of improved security requirements. With an increase in the world population and the advent of globalization, traditional ways of maintaining security using manual record-keeping (like signature verification, photo/identity card verification) are no longer viable. Like most other tasks, identity verification processes have also been delegated to computers. Unique digital signatures are now being provided to people by matching digitally acquired biometric templates against the pre-stored templates in large biometric databases.

The task of automatic identity management has acquired pivotal importance due to the fast expansion of these biometric databases. Biometric based applications can now be seen in use around us in our day-to-day lives. Ruling bodies in many countries across the planet have taken up the task of providing their citizens with biometric identifiers and maintain identity databases. These databases are then used at airports and other entry points to regulate public movement across borders and single out suspicious elements. USA, Brazil, Germany, United Kingdom, Iraq, Israel, Australia, New Zealand etc have already started issuing passports containing digitized biometric data like signature, photographs, iris information etc. Many other countries are leading down the same path to maintain digital records of its population and are in the process of issuing passports with embedded biometric data.

Biometric identifiers can now be seen in ATMs and some web applications as well. This indicates the possibility of developing other important biometric applications, which have been relatively unexplored so far, and is the focus of this thesis. The obvious problem in porting the authentication technique being used in the ATMs using servers to enhance security of web applications is the lack of a suitable biometric sensor attached to our computers. Installing most biometric sensors is expensive and impractical due to its size. Moreover, once installed they need to be handled with care. This is a serious drawback to the otherwise immense potential of biometric authentication in commercially available civilian applications. On the other hand, there has been a steep decline in the prices of high definition cameras and a reduction in camera size due to which digital camera has now become accessible to the common man. This coupled with its easy availability, cameras can provide an alternative to the traditional modes of capturing biometric modalities like scanners and specialized imaging devices. This is a welcome change since having a light capturing device



Figure 1.1: Various applications of unconstrained palmprint imaging

as opposed to a bulky design will enable us to take the biometric technology to the doorsteps of a large bracket of people who now have easy access to good quality cameras.

Towards this end, we propose a generic camera based biometric authentication system to recognize individuals in unconstrained scenarios. Most laptops and mobile devices have in-built cameras and in systems that do not have one, it is easy to attach a generic web camera. Hence, any multi-purpose normal camera attached to a computing device along with it constitute our recognition system and thus making it a practical solution (See Figure 1.1). However, data collected using such a flexible system working in a real world situation results in many image variations which is of significance to the recognition algorithm. The fine details of the image are lost with introduction of a variety of unwanted additional parameters like illumination variation, noise, pose transformations, contrast issues etc. All these parameters adversely affect the recognition process, reducing the system accuracy enormously. This is one of the reasons for the state of art algorithms insist on an ideal image capture modality which has uniform and constant illumination settings, fixed pose and a restricted environment with a high quality sensor. Therefore, we need to develop recognition algorithms that can robustly handle the high degree of variability introduced due to this unconstrained and unsupervised image capture.

To summarise, we focus on the problems of deploying biometric authentication for commercial civilian applications. For this, we propose using a generic web camera connected to a computing device like a laptop or a mobile for capturing the biometric modality instead of the bulky and expensive state of the art scanners. Camera based biometrics allows for easy to use system modality for a user by following a flexible image capturing process. This flexibility entails fluctuations in the image quality and results in loss of texture and other variations viz. illumination, noise, pose variations etc. These variations have an adverse effect on the recognition accuracy. To overcome this problem, we require new algorithms designed to handle weak templates and to make use of the information available to us in the best possible way.

In this thesis, we propose a camera based recognition system for palmprint authentication. We present the system design and develop novel algorithms for identifying templates exhibiting unconstrained motion of hand. The introduction to thesis is organized as follows: In this chapter, we present an overview of biometrics. This includes a basic introduction to biometrics, followed by a section explaining the biometric recognition process. We then describe the set of metrics used to evaluate the performance of a biometric recognition algorithm followed by the various biometric modalities in popular use with a discussion on the factors to be considered when choosing the modality to be used by a given system. The chapter concludes by giving a detailed motivation to

the problem chosen for this thesis, exact problem statement, and thesis contributions.

1.2 Biometrics - An Overview

Biometrics refers to methods for uniquely recognizing humans based upon one or more **physical** or **behavioral** traits.

- **Physiological** characteristics are related to the physical characteristics of the body. Examples include fingerprints, face , DNA, hand and palm geometry, iris feature, which has largely replaced retina, and odor/scent.
- **Behavioral** characteristics are related to the behavior of a person. Examples include gait, and voice. Some researchers[1] have coined the term behaviometrics for this class of biometrics.

In information technology, in particular, biometrics is used as a form of access control. Biometric technologies are becoming the foundation of highly secure identification and personal verification solutions. With an increase in level of security breaches and transaction frauds, the need for highly secure identification and personal verification technologies is becoming apparent.

In the current day and age of an internet savvy population, the world wide web has become an important medium to carry out social interchanges, monetary transactions and has emerged as a storehouse of personal information. According to a survey [2], an astonishing 1.7 billion users access the internet everyday as of the year 2009. While this staggering number showcases the success of the world wide web, it also subtly spells out a new genre of problems related to internet thefts and infringement of an individual's personal space. The term 'Personal security' has encompassed new dimensions in this era. In the light of increasing number of hacking attacks, passwords no longer guarantee an individual's privacy or safety of valuable information. It is possible to crack a password by using simple randomized algorithms that generate strings of different lengths with all possible combinations of the alphabet space. There are many softwares available in the market for the purpose of password cracking. Hence, there is a need for better ways to secure the financial and/or personal interchanges that we now carry out everyday on the internet.

Since the inception of world wide web in 1968 ([3]), its application has expanded tremendously. It is now being used as an entertainment center, a medium for exchanging information, a place for business transactions and as a new way of socializing. Web-based applications are more popular than ever. They generally work by assigning a user name and a password enabling one to access their "account". With the increase in the number of applications, people have more and more accounts to be managed. It is highly recommended to have different passwords for different accounts, so that, if the password for one account is compromised, others remain safe. Hence, there is an unnecessary task of remembering so many passwords. This can be especially tricky when carrying out financial transactions. In the face of these security threats, biometrics has emerged as a potential alternative solution to remembering passwords and carrying RFID cards. One does not need to worry about remembering different security combinations in a security model where access is granted on the basis of "Who you are?", instead of "What you remember?" (See Figure 1.2). This ease of use makes biometrics immensely convenient and handy. It can also help avoid fraudulent accounts where the same person is using more than one bank account by linking each account with a unique biometric id. This kind of a system is in use in the world famous amusement and theme

park, Walt Disney World situated in Florida, USA. At Walt Disney, biometric measurements are taken from the fingers of guests to ensure that the person's ticket is used by the same person from day to day ([4]).

Biometric-based authentication applications include workstation, network, and domain access, single sign-on, application log-on, data protection, remote access to resources, transaction security and Web security. Trust in these electronic transactions is essential to the healthy growth of the global economy. Utilized alone or integrated with other technologies such as smart cards, encryption keys and digital signatures, biometrics are set to pervade nearly all aspects of the economy and our daily lives. Utilizing biometrics for personal authentication is becoming convenient and considerably more accurate than current methods (such as the utilization of passwords or PINs). This is because biometrics links the event to a particular individual (a password or token may be used by someone other than the authorized user), is convenient (nothing to carry or remember), accurate (it provides for positive authentication), can provide an audit trail and is becoming socially acceptable and inexpensive [5].

In summary, biometrics is a robust authentication method since it cannot be stolen, forgotten, borrowed or hacked easily by standard password cracking softwares, and forging one is practically impossible. However, one advantage of passwords over biometrics is that they can be re-issued. If a token or a password is lost or stolen, it can be cancelled and replaced by a newer version. This is not naturally available in biometrics. If someones face is compromised from a database, they cannot cancel or reissue it. **Cancelable Biometrics** [6] is a way in which to incorporate protection and the replacement features into biometrics. Hence, Biometric technologies overcome the shortcomings of password and token based identification strategies while retaining their advantages. The strongest security measure is considered to be a mixture of the three (See Figure 1.2, [7]).

1.2.1 Classes of biometric systems

Biometric-based solutions are able to provide for confidential financial transactions and personal data privacy. The need for biometrics can be found in federal, state and local governments, in the military, and in commercial applications. Enterprise-wide network security infrastructures, government IDs, secure electronic banking, investing and other financial transactions, retail sales, law enforcement, and health and social services are already benefitting from these technologies. A natural question arises as to how does one determine which biometric modality or solution is applicable to the application in question. This is determined by studying the specific requirements of the application being used in the multi-dimensional parameter space created with various classes of biometrics as the independent axes. These biometric characteristics have been defined here. For a given biometric application, any physical or behavioral characteristic can be used as a biometric only if it rates well on the parameters stated below:

- **Universality:** This means that every person should have the characteristic.
- **Uniqueness:** This measures the capacity of the biometric to distinguish a person from all the others.
- **Permanence:** This states how well a biometric resists aging and others variations over time.
- **Collectability:** It refers to the ease of acquisition for measurement.

- **Performance:** This is a measure of the accuracy, speed and robustness of the technology used.
- **User Acceptability:** is the term given to the response generated by the biometric characteristic among the subjects who are to use the technology. It basically refers to the ease of use for the subject.
- **Circumvention:** refers to how easy it is to fool the system.

This can be easily seen by considering few specific applications of biometrics. As already described in 1.1, many countries now issue passports having digital signatures imbedded in them. These include photograph, ten fingerprints, handwriting sample etc. This helps the government regulate movement across borders. Since most people have passports, this has result in huge databases of individual digital signatures. Identity record management on a huge scale has also benefitted in tracking criminals and other wrong doers. Biometric databases are highly useful for forensic scientists as well. They now have better recognition softwares and easy access to huge databases at their disposal. This hastens the process of law enforcement. These are the applications that need stringent security measures. One wrong decision can incriminate an innocent man. It is for this reason that the biometric setup and the modalities used for these applications are required to be highly accurate with minimum False Positives. Hence, we would like to design a system that optimizes the operating point of such systems. This requires selecting a biometric modality (various physiological characteristics used for identification such as fingerprint, iris, retina etc are called modalities) that is highly accurate. Here, we give low priority to other characteristics of a biometric recognition system like user acceptance and comfort, cost of installation, maintenance charges etc (See Section 1.6). Hence, we choose precise modalities like iris, retina or fingerprints over modalities such as face, palmprint or hand geometry etc which are less accurate but are more user-friendly and easy to use.

In short, parameters such as false alarms, genuine rejects, user convenience, cost and computational requirements are given different weightage in a biometric recognition system, depending on the application. For example, in the applications described so far, we need a high level of security and hence, low False Positives(see Section 1.5). But there could be systems that are designed for everyday use by the normal public viz. authentication to get fixed ration. In this case, we don't mind if a couple of users have been falsely recognized (high False Positive). Our focus is on the fact that the actual enrolled users should never be turned down by the system. This means that we need to design a system with high True Positives. Another consideration in designing such a system is that since this system shall be used frequently by the enrolled subjects(to get ration regularly), they would not want to go through a lengthy and uncomfortable authentication process. Therefore, this kind of a system will use a biometric modality that is easy to capture and is quickly processed for example, face, hand, palm etc.

Hence, we can conclude that designing a biometric system requires an in-depth analysis of the system requirements first. It should then be optimized with respect to the many parameters that constitute as various biometric classes. Due to its versatile nature, biometrics can be used for a wide spectrum of applications. The same model or algorithm cannot fit two different scenarios and hence cannot be directly ported. Some are applicable in a closed laboratory environment while others solve issues relevant in the real world in our day-to-day life. The technical challenges faced in each of these application areas are quiet diverse. The potential for research lies in not only

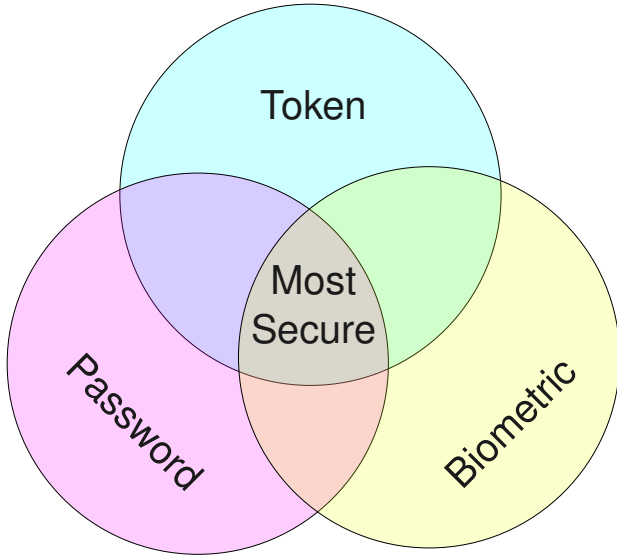


Figure 1.2: Ways of Authentication

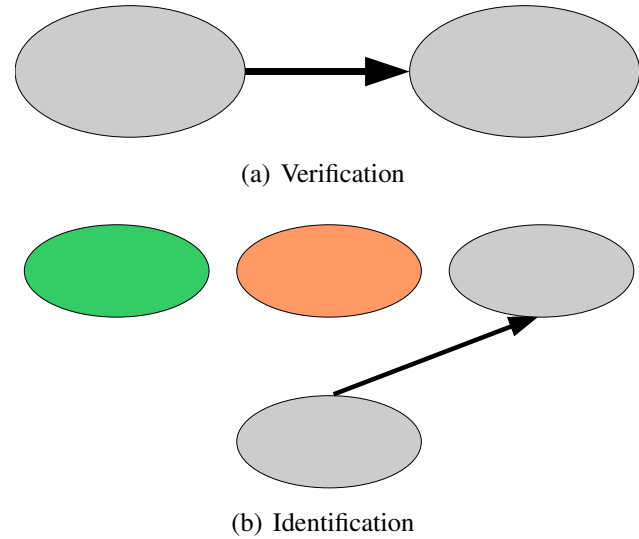


Figure 1.3: Modes of Operation

developing new biometric modalities(viz. ear, foot, tongue based recognition etc.); but also to develop new and efficient algorithms to face some of the key technical challenges in each of these modalities when used in a specific application. We describe these scenarios in more detail in this section.

The practice of maintaining digital biometric records being followed in some countries like USA, UK, Germany etc will, in future, have many advantages apart from fulfilling security requirements. These databases can be used for counting unique votes cast in a region and to manage a complete profile of a user including bank account numbers, PAN card numbers etc. People need not remember long strings of these bank and PAN card numbers as once identified correctly, the system can provide all the required information. Biometrics based attendance monitoring systems in schools, colleges and workplaces can be put in place to ensure no proxies or fake records resulting in a highly productive workforce. One could also monitor various trends prevalent in a population segment and provide an understanding of how people operate in their day to day life. For example, a grocery store maintaining biometric records can establish the age groups that shops more for a particular item, resulting in better customer satisfaction. The maintained records can help keep track of the individuals availing various schemes provided by the government and hence ensuring that there are no fraudulent claims and that the right people are reaping the benefits of the scheme as in the services envisaged under the unique ID project.

To summarise, biometrics algorithms provide fast and effective solutions for a large variety of security issues faced by us in our day-to-day lives. It is a complex optimization in the space of many parameters, namely, installation cost, ease of use, user acceptability, level of security needed, need for supervision, maintenance cost etc. Forensic applications and airport security both require tight bounds on identification, but airport security system should also be efficient in real-time, which is not an critical for a forensic application. Similarly, various other applications need to optimized depending upon the requirement and its target audience.

1.3 Biometric Recognition Process

The first time an individual uses a biometric system is called an **enrollment**. During the enrollment, biometric information from an individual is stored (See Figure 1.3). In subsequent uses, biometric information is captured and compared with the information stored at the time of enrollment. Note that it is crucial that storage and retrieval of such systems themselves be secure if the biometric system is to be robust. The first block (sensor) is the interface between the real world and the system; it has to acquire all the necessary data. Most of the times it is an image acquisition system, but it can change according to the characteristics desired. The second block performs all the necessary pre-processing: it has to remove artifacts from the sensor, to enhance the input (e.g. removing background noise), use transformations such as normalization and scaling. In the third block, the required features are extracted. This step is an important step as the quality of features have a significant effect on the accuracy of a biometric system. A vector of numbers or an image with particular properties is used to create a template. A template is a synthesis of the relevant characteristics extracted from the source. Elements of the biometric measurement that are not used in the comparison algorithm are discarded in the template to reduce the file size and to protect any personal information of the enrollee.

If enrollment is being performed the template is simply stored somewhere (on a card or within a database or both). If a matching phase is being performed, the obtained template is passed to a matcher that compares it with other existing templates, estimating the distance between them using any algorithm (e.g. Hamming distance [8]). The matching program will analyze the template with the input. This will then be output for any specified use or purpose (e.g. entrance in a restricted area).

1.3.1 Biometrics: Modes of Operation

There are two ways in which a Biometric System can authenticate individuals (See Figure 1.3,[7]).

Verification

- A **one-to-one** comparison of a captured biometric with a stores template to verify that the individual is who he/she claims to be. It can be carried out in conjunction with a smart card, user name or ID number.

Identification

- A **one-to-many** comparison of the captured biometric against a database in attempt to identify an unknown individual. The identification only succeeds in identifying the individual if the comparison of the biometric sample to a template in the database falls within a previously set threshold.

1.4 Performance Metrics

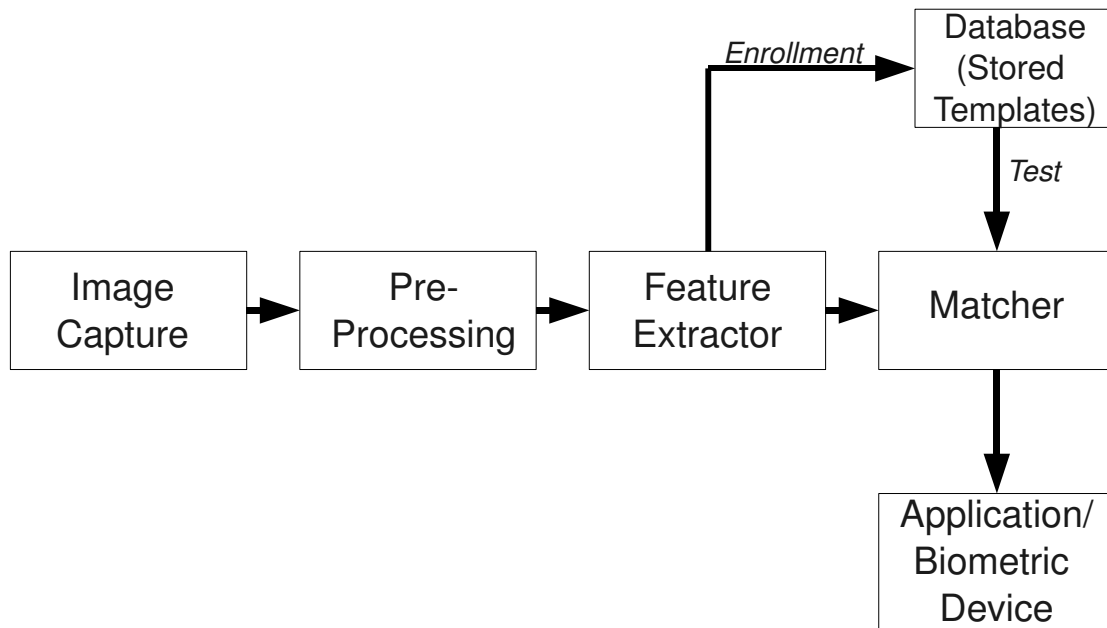
The performance of a biometric system is estimated using the classification model used in the signal detection theory. ROC (Receiver Operating Characteristics) is a visual representation for

a better understanding of the recognition accuracy, and EER (Equal Error Rate) is used as an approximated absolute measure of accuracy. To comprehend these terms, we first need to understand the underlying classification model.

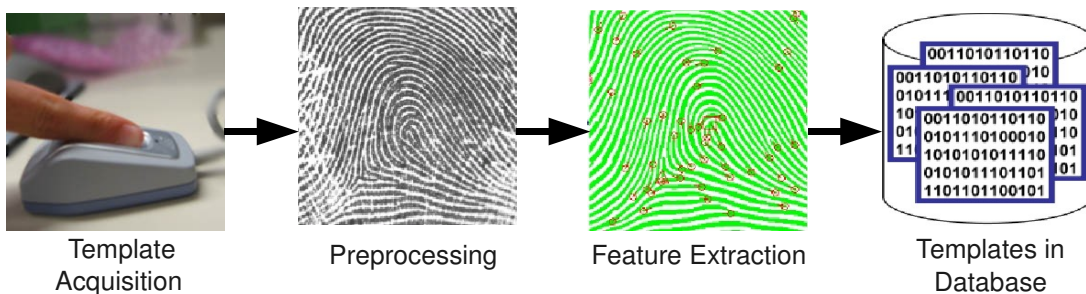
1.4.1 The Basic Model

A classification model (classifier or diagnosis) is a mapping of instances into a certain class/group. The classifier or diagnosis result can be a real value (continuous output) in which the classifier boundary between classes must be determined by a threshold value. For instance, to determine whether a person has hypertension based on blood pressure measure. Result can also be in a discrete class label indicating one of the classes.

Let us consider a two-class prediction problem (binary classification), in which the outcomes



(a) The basic block diagram of a Biometric Recognition System



(b) The process of enrollment taking fingerprint as an example.

Figure 1.4: A Biometric Recognition System.

		actual value	
		p	n
prediction outcome	p'	True Positive	False Positive
	n'	False Negative	True Negative
total		P	N

Table 1.1: Confusion Matrix

are labeled either as positive (p) or negative (n) class. There are four possible outcomes from a binary classifier. If the outcome from a prediction is p and the actual value is also p , then it is called a true positive (TP); however if the actual value is n then it is said to be a false positive (FP). Conversely, a true negative has occurred when both the prediction outcome and the actual value are n , and false negative is when the prediction outcome is n while the actual value is p .

To get an appropriate example of a real-world problem, consider a diagnostic test that seeks to determine whether a person has a certain disease. A false positive in this case occurs when the person tests positive, but actually does not have the disease. A false negative, on the other hand, occurs when the person tests negative, suggesting they are healthy, when they actually do have the disease.

Let us define an experiment from P positive instances and N negative instances. The four outcomes can be formulated in a 2×2 contingency table or confusion matrix, as shown in Table 1.1.

Based on this contingency table, several performance evaluation metrics can be derived:

- **True Positive (TP)** - If both, the outcome and the actual value of a prediction are positive, then its said to be a True Positive. This is the equivalent of a “hit” in Signal Detection Theory.
- **True Negative (TN)** - If both, the outcome and the actual value of a prediction are negative, then its called a True Negative. This is the equivalent of a “correct rejection” Signal Detection Theory.

- **False Positive (FP)** - If the actual value is negative and the outcome is positive, then its a False Positive. This is the equivalent of a “*false alarm*” in Signal Detection Theory.
- **False Negative (FN)** - If the actual value is positive and the outcome is negative, then it is said to be a False Negative. This is equivalent to a “*miss*” in Signal Detection Theory.

Please note that in the following text,

$$\begin{aligned} P &= \text{Total number of positive/correct samples, and} \\ N &= \text{Total number of negative/incorrect samples .} \end{aligned}$$

- **Sensitivity or true positive rate (TPR)** - It is the probability with which the system correctly matches a template with an existing pattern in the database. This concept is equivalent to “*hit rate*” or “*recall*”.

$$TPR = TP \div P = TP \div (TP + FN), \quad (1.1)$$

- **False Positive Rate (FPR)** - It is the probability with which the system falsely or incorrectly matches a template with an existing pattern in the database. This is the equivalent of “*false alarm rate*” or “*fall-out*”.

$$FPR = FP \div N = FP \div (FP + TN), \quad (1.2)$$

- **Accuracy (ACC)** - This refers to the probability with which s template(or sample) is correctly identified.

$$ACC = (TP + TN) \div (P + N), \quad (1.3)$$

- **Specificity(SPC) or True Negative Rate** - A specificity of 100% means that no positives are erroneously tagged.

$$SPC = TN \div N = TN \div (FP + TN) = 1 - FPR, \quad (1.4)$$

Now, we have all the required metrics in place, we can look into methods of rating the performance of biometric systems. The most commonly used terms are the ROC (Receiver Operating Characteristics) and EER (Equal Error Rate).

1.4.2 ROC Space

We model a Biometric authentication problem as a two-class problem for both verification and identification tasks. In verification, if the given template matches with the claimed identity, the outcome is *true* (called a *genuine* match), and *false* (called an *imposter*) otherwise. In identification, if the query template matches with a person in the database, the system returns the *user id* (termed as a *genuine* result) of the matched entry, and returns *false* (called an *imposter*)if no match is found(See Figure 1.7). In a two-class classification problem, decision is taken using

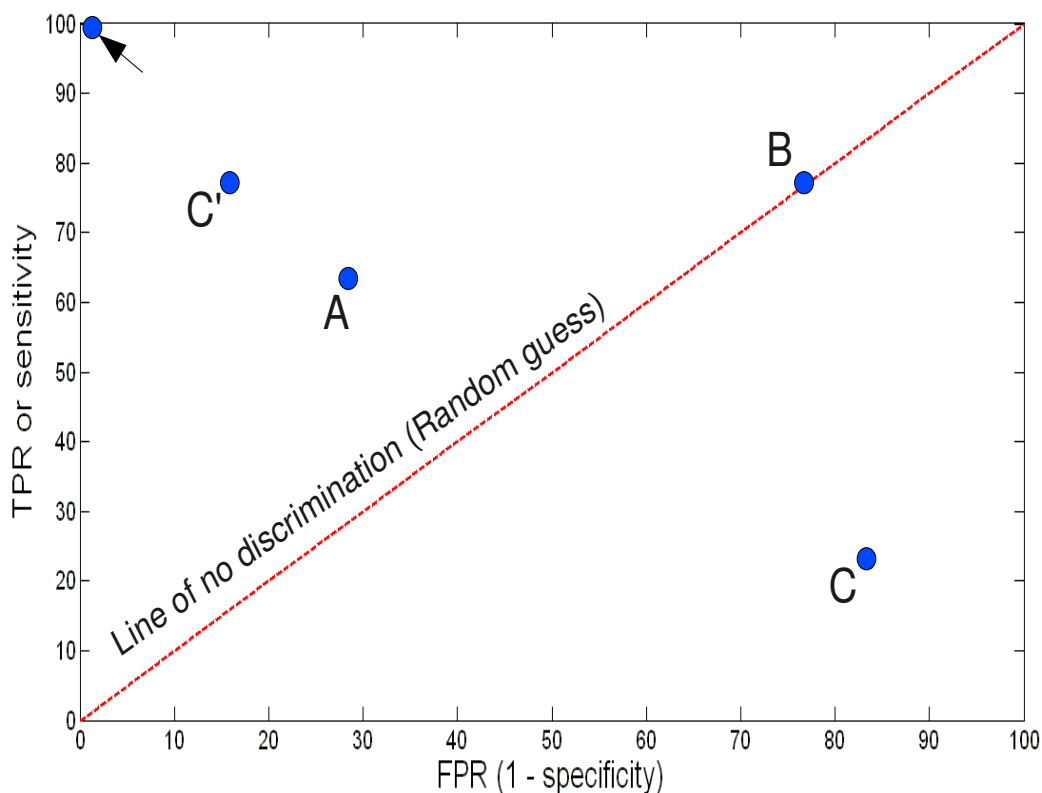


Figure 1.5: ROC space and plots of the four prediction examples

thresholding on the feature vector probability. In signal detection theory, a receiver operating characteristic (ROC), or simply ROC curve, is a graphical plot of the *sensitivity* vs. (*specificity*) for a binary classifier system as its discrimination threshold is varied. The ROC can also be represented equivalently by plotting the fraction of true positives ($TPR = \text{true positive rate}$) vs. the fraction of false positives ($FPR = \text{false positive rate}$). Also known as a Relative Operating Characteristic curve, because it is a comparison of two operating characteristics (TPR & FPR) as the criterion changes [9].

TPR determines a classifier or a diagnostic test performance on classifying positive instances correctly among all positive samples available during the test. FPR , on the other hand, defines how many incorrect positive results occur among all negative samples available during the test [10]. ROC curves are very useful for the evaluation of machine learning techniques. The first application of ROC in machine learning was by Spackman who demonstrated the value of ROC curves in comparing and evaluating different classification algorithms [11].

A ROC space is defined by FPR and TPR as x and y axes respectively, which depicts relative trade-offs between TP , true positive (benefits) and FP false positive (costs). Since TPR is equivalent with sensitivity and FPR is equal to $1 - \text{specificity}$, the ROC graph is sometimes called the *sensitivity* vs. ($1 - \text{specificity}$) plot. Each outcome or one instance of a confusion matrix represents one point in the ROC space.

The best possible prediction method would yield a point in the upper left corner or coordinate

(0, 1) of the ROC space, representing 100% *sensitivity* (no false negatives) and 100% *specificity* (no false positives) (See Figure 1.5). The (0, 1) point is also called a *perfect classification*. A completely random guess would give a point along a diagonal line (the so-called line of no-discrimination) from the left bottom to the top right corners. An intuitive example of random guessing is a decision by flipping coins (head or tail).

The diagonal line divides the ROC space in areas of good or bad classification/diagnostic. Points above the diagonal line indicate good classification results, while points below the line indicate wrong results (although the prediction method can be simply inverted to get points above the line). Let us look into four prediction results from 100 positive and 100 negative instances as depicted by Table 1.2.

A			B			C			C'		
TP=63	FP=28	91	TP=77	FP=77	154	TP=24	FP=88	112	TP=76	FP=12	88
FN=37	TN=72	109	FN=23	TN=23	46	FN=76	TN=12	88	FN=24	TN=88	112
100	100	200	100	100	200	100	100	200	100	100	200
TPR = 0.63			TPR = 0.77			TPR = 0.24			TPR = 0.76		
FPR = 0.28			FPR = 0.77			FPR = 0.88			FPR = 0.12		
ACC = 0.68			ACC = 0.50			ACC = 0.18			ACC = 0.82		

Table 1.2: Confusion matrices corresponding to different classification or prediction systems.

Plots of the four results above in the ROC space are given in the Figure 1.5. The result *A* clearly shows the best among *A*, *B*, and *C*. The result *B* lies on the random guess line (the diagonal line), and it can be seen in the table that the accuracy of *B* is 50%. However, when the matrix *C* is mirrored about the central horizontal line, as seen in *C'*, the result is even better than *A*. Since this mirrored *C* method or test simply reverses the predictions of whatever method or test produced the *C* contingency table, the *C'* method has positive predictive power simply by reversing all of its decisions. When the *C* method predicts *p* or *n*, the *C'* method would predict *n* or *p*, respectively. In this manner, the *C'* test would perform the best. While the closer a result from a contingency table is to the upper left corner the better it predicts, the distance from the random guess line in either direction is the best indicator of how much predictive power a method has, albeit, if it is below the line, all of its predictions including its more often wrong predictions must be reversed in order to utilize the method's power.

Curves in ROC Space

Discrete classifiers, such as decision tree or rule set, yield numerical values or binary label. Applied to a set of instances, such classifiers yield a single point in the ROC space. Other classifiers, such as naive Bayesian and neural network, produce probability values representing the degree to which class the instance belongs to. For these methods, setting a threshold value will determine a point in the ROC space. For instance, if probability values below or equal to a threshold value of 0.8

are sent to the positive class, and other values are assigned to the negative class, then a confusion matrix can be calculated. Plotting the ROC point for each possible threshold value results in a curve.

Further Interpretations

Sometimes, the ROC is used to generate a summary statistic. Three common versions are:

- The intercept of the ROC curve with the line at 90 degrees to the no-discrimination line gives the values of TPR and FPR for a particular threshold value t_0 (See Figure 1.6).
- The area between the ROC curve and the no-discrimination line. A greater area indicates a better recognition system.
- The area under the ROC curve, or “ AUC ”, or A' [12]. The AUC is equal to the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.
- The distance, d' , between the mean of the distribution of activity in the system under noise-alone conditions and its distribution under signal-alone conditions, divided by their standard deviation, under the assumption that both these distributions are normal with the same standard deviation. Under these assumptions, it can be proved that the shape of the ROC depends only on d' .

However, any attempt to summarize the ROC curve into a single number loses information about the pattern of tradeoffs of the particular discriminator algorithm.

The machine learning community most often uses the ROC AUC statistic for model comparison [11]. This measure can be interpreted as the probability that when we randomly pick one positive and one negative example, the classifier will assign a higher score to the positive example than to the negative. In engineering, the area between the ROC curve and the no-discrimination line is often preferred, because of its useful mathematical properties as a non-parametric statistic. This area is often simply known as the discrimination. In psychophysics, d' is the most commonly used measure.

Sometimes it can be more useful to look at a specific region of the ROC Curve rather than at the whole curve. It is possible to compute partial AUC [13]. For example, one could focus on the region of the curve with low false positive rate (FPR), which is often of prime interest for population screening tests [14]. As the sensitivity of the biometric device increases, the FAR decreases but the FRR increases. Another common approach for classification problems in which $P \ll N$ (common in bioinformatics and biometrics applications) is to use a logarithmic scale for the x-axis [15].

1.4.3 Equal Error Rate (EER)

The **Equal Error Rate** or the **Crossover Error Rate** is the rate at which both the accept and reject errors are equal. The value of the EER can be easily obtained from the ROC curve. The EER is a quick way to compare the accuracy of devices with different ROC curves. In general,

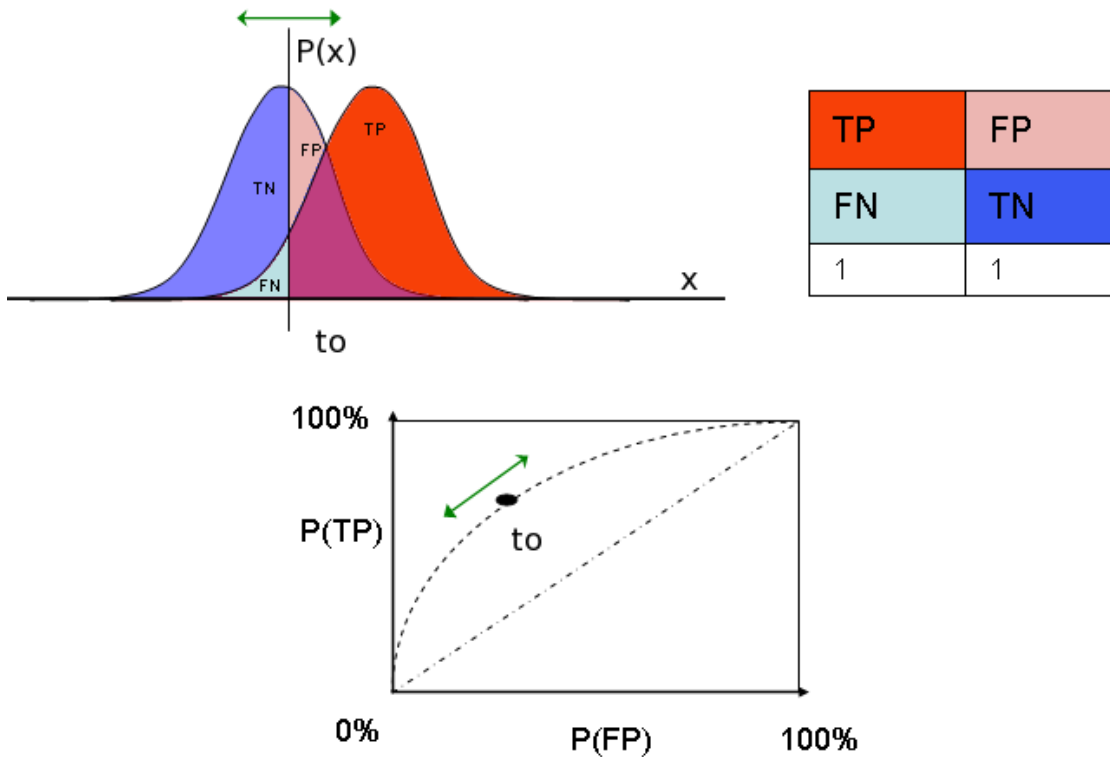


Figure 1.6: ROC curve for a univariate two class probability distribution. Each point on a ROC curve corresponds to a particular confusion matrix at a particular threshold value t_0 .

the device with the lowest $EEER$ is most accurate. Obtained from the ROC plot by taking the point where FAR and FRR have the same value. The lower the $EEER$, the more accurate the system is considered to be.

1.4.4 Others

The ROC curve and the $EEER$ values are useful in rating the efficiency of the system. There are some other aspects of a biometric system, however, that cannot be defined by these two values. They are explained below:

- **Failure to Enroll Rate (FTE or FER)** The rate at which attempts to create a template from an input is unsuccessful. This is most commonly caused by low quality inputs or in cases when the user does not present the biometric in accordance with the prescribed guidelines.
- **Failure to Capture Rate (FTC)** Within automatic systems, the probability that the system fails to detect a biometric input when presented correctly.

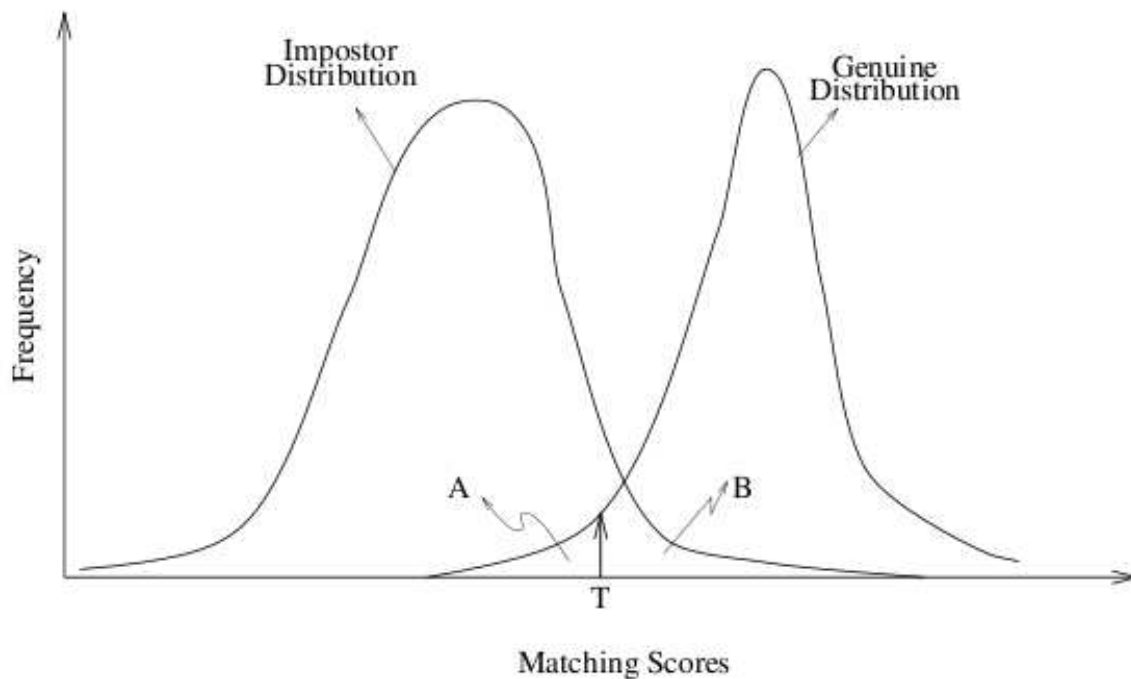


Figure 1.7: Probability distribution of the two class data points in Biometrics, namely the genuine and the imposter matching scores.

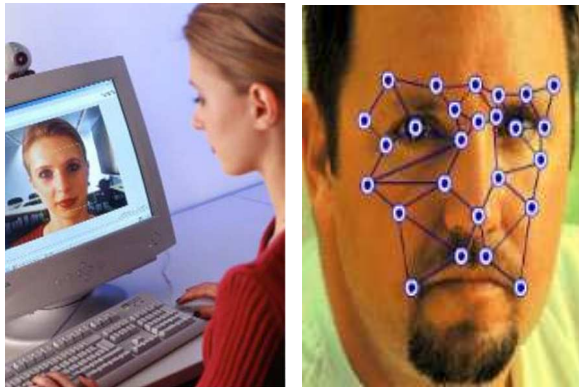
- **Template Capacity** This refers to the maximum number of sets of data which can be stored in the system.

1.5 An overview of Biometric Technologies

A number of physiological and behavioral features have been proposed and proved to be efficient biometric characteristics. But, every biometric characteristic has its limitations, and no single biometric is expected to completely distinguish every individual. Therefore, a lot of research has been done to establish the strengths and weaknesses of various biometric features. A short summary of the advantages and shortcomings of various biometric modalities has been provided below.

1.5.1 Fingerprints

Fingerprints are the flow-like ridges present on human fingers [16]. They are believed to be unique for each person (and each finger). This is one of the most mature biometric technology, being used in forensic divisions for criminal investigation. The features present on the fingerprint used for distinguishing are the ridge patterns and the fingerprint minutiae, which are the features resulting mainly from ridge endings and bifurcations (Figure 1.9(b)). This hints at the level of detail that is required to be captured in an image to be able to distinguish fingerprints.



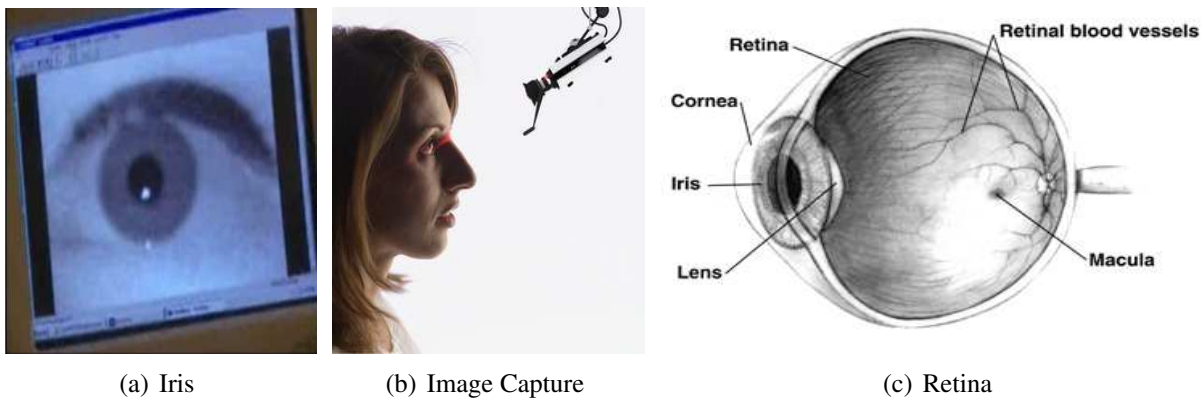
(a) Image Acquisition (b) Facial Features

Figure 1.8: Facial Recognition



(a) Fingerprint Scan (b) Features

Figure 1.9: Fingerprint Based Recognition



(a) Iris (b) Image Capture (c) Retina

Figure 1.10: Iris and Retina Recognition.

A fingerprint is captured in two ways: a) by scanning an inked impression of a finger (Figure 1.9(a)), or b) using a live scanner. Typical systems that are in use now use a resolution of 500 DPI. Although it is possible to obtain images of better quality now. But a touch-based scanner is a source of hygiene based concerns and maintenance problems. These issues rate high in determining the usability of a biometric system. Hence, there is a shift towards developing algorithms that can use digital camera based imaging technologies [17]. Once the fingerprint has been captured, the image is then used for either feature extraction or the gray scale values can be used as a feature vector itself. The results obtained by integrating more than one fingerprint are phenomenal as it greatly enhances the security of the system.

1.5.2 Face

Facial features is what humans recognize each other by. It comes naturally to human beings. But it is not that straight forward for computers. A face recognition algorithm uses the following two methods to recognize human faces: a) *Transform Approach*: Project the face image into a set of orthonormal basis vectors([18],[19]), b) *Feature Based*: By selecting the following facial features: the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw(See Figure 1.8(b)).

The database is then searched for an image having the matching feature set [20]. It is not one of the most reliable and efficient biometrics around, but its greatest usability as a biometric lies in its ready acceptance by the public. Moreover, subjects don't need aid to present face to a system and is also helpful for the police in recognizing criminals in a crowd etc [21]. It is for this reason that face recognition has been readily installed in airport facilities and other public places.

But the advantage of user flexibility is also the main limitation of a face recognition system. On giving subjects the freedom to pose at free will; the system is easily affected by factors such as sunglasses, different hairstyles, the absence/presence of beards, face covering hair, hats etc. To add to these, there are pose variations and lighting issues. But because of the obvious advantages, research is on to diminish the limiting factors of face recognition on various fronts. Some of the algorithms are now able to outperform human participants in recognizing faces and could uniquely identify identical twins [22]. Hence, due to ongoing research, it is now possible to design face recognition algorithms that use camera based image capture devices under variable lighting conditions as well(Figure 1.8(a) [18]).

1.5.3 Iris

Iris recognition is a method for personal identification which is considered to be highly accurate. The iris pattern is created as a result of embryonic genetic gestation just like fingerprints(Figure 1.10(a)). It is captured using a camera having an infrared sensor; the person has to look into the camera while the image is being captured. Hence, it is free from the usual hygiene concerns as it employs a contact-less imaging system(See Figure 1.10(b)). It is also thought to have the highest *template longevity*, as, barring trauma, a single enrollment can last a lifetime [23]. Unlike face and hand, the degree of variations associated with iris in terms of shape are negligible since it can only dilate and contract; which makes its structure predictable.

The only disadvantage associated with iris is that it cannot be captured from a distance of more than a few meters. Moreover, iris capture needs supervision and user cooperation as iris needs to be shown at a predetermined distance from the camera and in a way such that the central region is clearly visible in the image. As with other photographic biometric technologies, iris recognition is susceptible to poor image quality, with associated failure to enroll rates [24].

1.5.4 Retina

Retina is a thin tissue having vascular patterns located at the posterior part of the eye(See Figure 1.10(c)). It is believed to be a unique biometric because of the the complex structure of the capillaries which supply retina with blood[25]. To capture this pattern, a person has to look into an eyepiece. A low intensity unperceivable ray of infrared light is then sent to light up the retina and capture the image. Since in this image capture process, the person has to be in contact with the apparatus and expose a very sensitive part of the body, namely eyes, it ranks low on user acceptance. Please note that this is in contrast to the iris scan which can be taken at a distance of few meters. It is also possible to examine a person's retina and predict medical conditions of diabetes, glaucoma or retinal degenerative disorders; which is sensitive personal information. This is another reason because of which people are apprehensive about retinal identification methods even though it is believed by some to be the most precise and reliable biometric [26].

1.5.5 Gait Recognition

Gait Recognition is referred to the technique used to identify a person by his/her walking style(See Figure 1.13). It is a behavioral biometric modality which recognizes individuals by using the principles of kinematics in the spatio-temporal domain. A person is usually asked to move on a treadmill or a specially designed track connected to a computer. A small video footage is then made to capture the complete movement [27]. Then attributes such as step length, step width, walking speed, joint rotation of the hip, knee and ankle, mean joint angles of the hip/knee/angle etc. are recorded to make a feature vector.

Gait is a weak biometric since walking patterns of individuals tend to differ with age, physical injury or physical conditions(like pregnancy, fatigue etc). There are no commercial applications available using gait recognition currently but it is expected to rate high on user acceptance since like face recognition it can be captured using only a camera. Since it is very easy to obtain data having lots of people walking in surveillance videos, gait recognition might prove very useful in identifying suspects from a distance. Hence, research is ongoing in this area to identify gait in the presence of lighting variations, occlusions etc. to make the technology more robust for use even in crowded scenarios as gait is an useful biometric add on when combined with overall appearance and facial features of a person[28].

1.5.6 Keystroke Dynamics

Keystroke Dynamics is the detailed timing information that describes exactly when each key was depressed and when it was released as a person is typing at a computer keyboard [29]. It is another behavioral biometric modality which studies the manner and rhythm of the typing style of an individual. The common differences that occur in individual typing styles are these: a) A person might use either of right or left *shift* key, b) Capital letters can be generated using a *Caps Lock* key or by using *shift + alphabet*, c) speed of typing, d) difference in typing speed when typing characters as opposed to typing numbers(or symbols), e) frequency of using the *backspace* or *delete* key etc (Figure 1.11).

It is a weak biometric which means that it has limited differential capacity but rates high in user acceptability which makes it ideal for use in a civilian application. One of the particularly interesting applications is to track a person who is not authorized to use a system but is doing so illegally. Since, it records and analyzes data in the temporal domain continuously(and not just at the start up time), a silent alarm can be sent to the system administrator or any other concerned party to alert them of the security breach. The shortcoming of keystroke dynamics is the variation incurred due to change in the mood or state of the person. People will have different typing speeds and styles when they are tired or inebriated as compared to when they are feeling normal. There is a need to develop robust systems to take these variations into account [30].

1.5.7 Handwriting Based Recognition

Handwriting analysis is also a part of the study of behavioral biometrics. Signature matching on documents of importance is one of the commonly heard of applications of this matching technique. The signature size is first normalized and then broken down into small parts called the strokes which are then compared against other users to uniquely identify an individual's style. As other



Figure 1.11: Keystroke Based Recognition



Figure 1.12: Signature



Figure 1.13: Gait

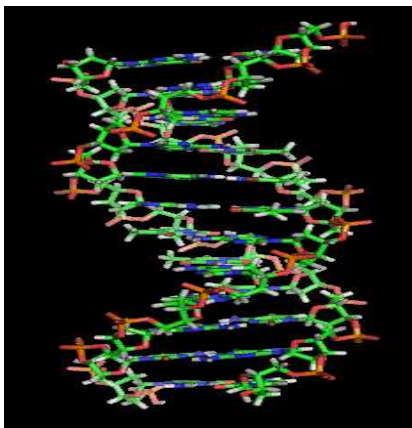
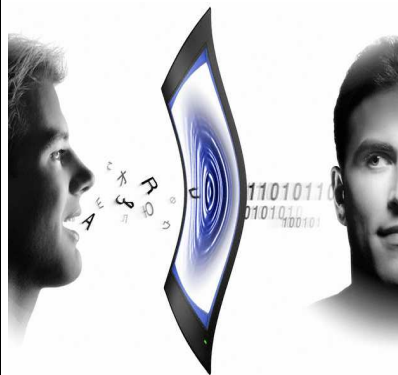
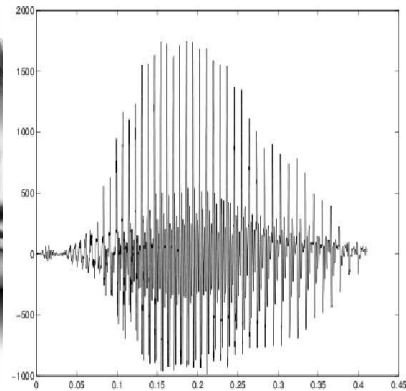


Figure 1.14: DNA Profiling



(a)



(b) Time vs. Signal plot for the word 'seven'

Figure 1.15: Speaker Recognition

behavioral biometric modalities, it is also a weak biometric and depends on the state of mind of a person and exhibits a wide array of variations which poses many challenges in the design of an automatic signature recognizer.

Ongoing research focuses on developing automatic handwriting recognition systems that can be classified in the following two categories: a) Off-line handwriting recognition, where a subject is asked to write some text which is then used to identify individuals, and b) Online handwriting recognition, where the system tries to recognize subjects as they are writing[31]. The latter requires a touch-sensitive computer screen or a stylus (Figure 1.12), whereas off-line recognition methods can be used to identify individuals using scanned documents as well. Other interesting application that use handwriting recognition is postal-address interpretation and bank-cheque processing.

1.5.8 Voice Recognition

Speaker recognition is the computing task of validating a user's claimed identity using characteristics extracted from their voices. There is a difference between *speaker recognition* (recognizing

who is speaking) and *speech recognition* (recognizing what is being said). These two terms are frequently confused, as is voice recognition. Voice recognition is a combination of the two where it uses learned aspects of a speaker's voice to determine what is being said - such a system cannot recognize speech from random speakers very accurately, but it can reach high accuracy for individual voices it has been trained with [32].

The difference in the acoustic features of individuals is used as a discriminating characteristic. These acoustic patterns reflect both anatomy (e.g., size and shape of the throat and mouth) and learned behavioral patterns (e.g., voice pitch, speaking style) (Figure 1.15). The biometric is captured using an audio recorder in a sound proof room. Methods used to cancel noise may affect the system adversely if not tuned to the environment. There may be a variation in the data samples depending on the mood of the person providing the data. It rates high on user acceptability and many commercial softwares using voice recognition can be easily seen. Automatic Speaker Recognition is being actively researched upon [33].

1.5.9 Genetic Fingerprinting

Genetic fingerprinting (or DNA profiling) is the science of identifying individuals based on their DNA (Figure 1.14). DNA samples can be obtained from saliva, blood, skin samples from a crime scene. DNA profiling is done by comparing the lengths of variable sections of repetitive DNA, such as short tandem repeats and mini-satellites. DNA is the most reliable way known to identify individuals and is widely used in Forensics to solve crimes. It suffers with certain shortcomings though. It is very easy to steal DNA in the form of saliva and blood samples and hence enabling identity theft. Also, twins have the same genetic sequence of DNA. Because of its importance as a reliable evidence in a court of law, user acceptability for use in civilian applications is very low.

1.5.10 Hand Based Recognition

Hand Geometry

Hand geometry is a biometric that identifies users by the shape of their hands. Hand geometry readers measure a user's hand along many dimensions and compare those measurements to measurements stored in the database. To capture hand image, a dedicated setup having a platform to support the user's hand is used. Pegs are put to ensure fixed position of the hand (See Figure 1.16(a)). These pegs also help in locating the important landmarks on the hand. Algorithms to detect landmarks and contour without the help of pegs have also been developed [34]. The feature vector includes the length of the finger, palm and the hand, width of the fingers and hand at different points to capture the variation of geometry across the contour of the hand, and the thickness of the palm (See Figure 1.16(e)). This biometric has high user acceptance.

Palm Vein

This technology uses an infrared scan to examine the veins in an individual's palm (Figure 1.16(d)). Palm-vein patterns are unique to individuals, even individuals that share the same DNA. In fact, an individual user will have unique vein patterns even between a left hand and a right hand. People

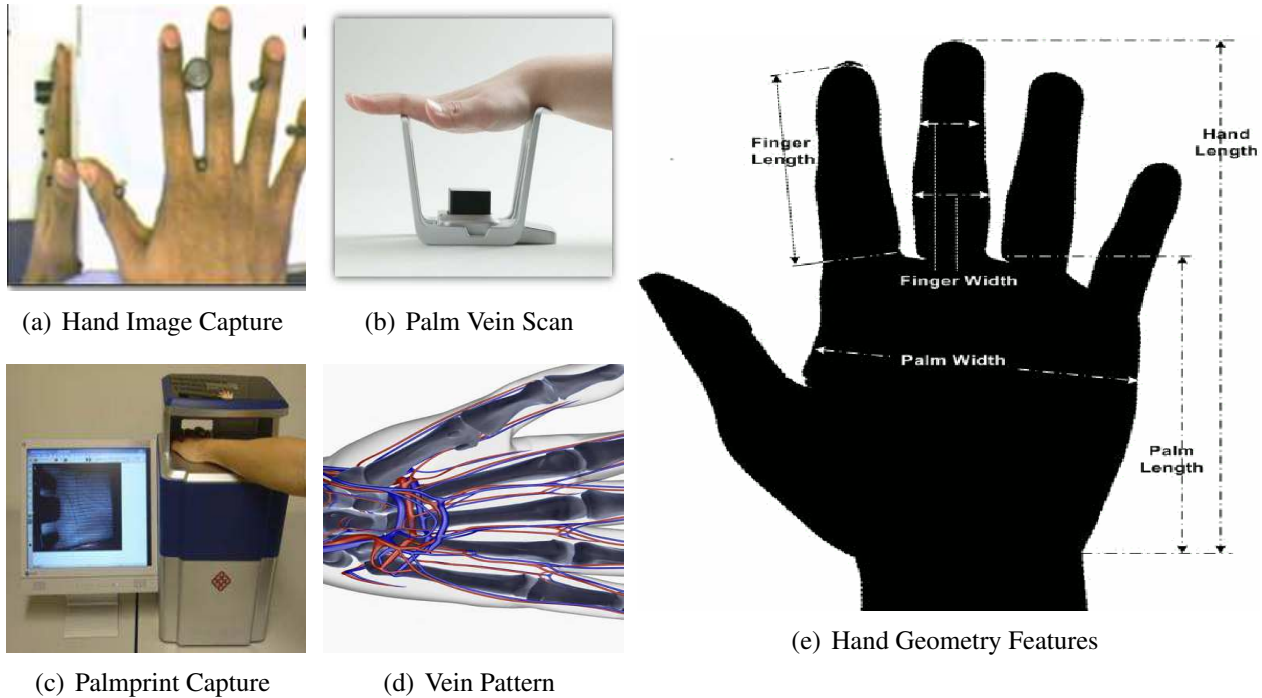


Figure 1.16: Hand Based Biometric Recognition

have high confidence in this technology because of the difficulty of “spoofing” someone’s unique vein pattern. This biometric has been successfully installed in ATMs [35].

Palmprint

Palm has unique distinguishing line patterns which can be used to identify people uniquely. It is a physiological biometric. The setup it uses is shown in Figure 1.16(a). A user is asked to put his/her hand in an enclosed unit having fixed illumination and a platform that restricts hand motion[8]. It may have pegs to further fix the position of the hand. Scanners or cameras can then be used to capture the image. Palmprint has high user acceptability with few hygiene based concerns characteristic of all contact based biometric. But there has been a shift towards developing contact-less recognition systems (2.1.1).

1.6 Choosing a Biometric Modality

The choice of a particular biometric modality depends heavily on the application for which it is being used. The factors to be considered are these:

- Whether the biometric is to be used in an authentication mode or a verification mode? If the biometric is to be used in an identification role, it should have fast computability and be more unique in general.

- Whether the security requirements of the system using the biometric are stringent. For example in cases of crime scene investigation and use in Forensic Labs, it is more important to have negligible false positives as we should not incriminate an innocent person. Whereas in a civilian application where a biometric is being used for access control, we can tolerate some false positives as the focus is on not denying access to the right person.
- Whether it is possible to supervise the biometric capture. In some remote places or in case of installation in an ATM device, it is better to use an automatic recognition device having a capturing device which is more intuitive to use.
- Whether the user is going to be co-operative. In a case where a huge number of attempts at forgery are expected, using a strong physiological characteristic such as retina, iris etc. is more reliable since it cannot be changed.
- Duration of time for which the biometric is to be used. If the application is going to be used for a long time, then it is advisable to use a biometric which is relatively more stable with time viz. iris pattern.
- Computational requirements associated with a particular biometric. This might vary with the change in availability of more sophisticated capturing devices.
- User acceptability. This is an abstract metric which depends on the cultural and social norms of the place at which the biometric is going to be used. People should be comfortable using a biometric. This is specially important to consider when using a biometric for commercial applications. Also, people are generally skeptical to give the more secure biometric modalities like fingerprints, iris and retina patterns at their local supermarkets for identification.
- The distance at the which the person has to be identified. This is an important feature in case of wanting to track an offender on loose. Biometric modalities capable of identifying individuals using features captured at a low resolution are preferred here.

Depending on the answers to the above questions, an organization may select which biometric modality is the most applicable to them. The impact of EER on different applications can be seen in Figure 1.17.

1.7 Introduction to Thesis

We now discuss the more specific concepts in some detail. In this section, we provide an introduction to the problem addressed in this thesis. Section 1.8 provides a general motivation towards the problem domain. Section 1.9 provides the problem statement followed by thesis contributions in the Section 1.10. The chapter ends with a thesis overview in 1.11.

1.8 Towards Camera Based Biometric Data Capture

A biometric recognition system has two basic steps: i) **Template Enrollment** and ii) **Matching of the new acquired template with the enrolled template**. Both the steps comprise of an

Image Capture module to capture the biometric data for enrollment and template matching, respectively. The Biometric capture module generally uses scanners or sensors for robust detection and is designed specific to the biometric modality being considered. This helps in taking special precautions to restrict the pose and illumination variations for a particular biometric. These factors, if left unchecked, greatly affect the recognition accuracy of the system by increasing the intra class variability. Hence, the state of the art biometric recognition algorithms use dedicated image capturing devices to get high quality scanned/digital images of the biometric modality.

Some of the dedicated image capturing modalities are: a) *Fingerprint Scanner*. The required finger is placed on the scanner which captures a detailed image of the finger (Figure 1.9(a)). b) *Iris/Retina image capture*: A specially designed camera having a small light source captures the image while the user is asked to stay still for a short period of time. (Figure 1.10(b)) c) *Face Recognition*: The user poses in front of a camera that captures the facial image. d) *Online Handwriting Scanners*. User are asked to write on a dedicated scanner connected to a PC (Figure 1.12). e) *Hand Geometry Capture Device*. The capturing setup has pegs on a fixed platform where the user places his/her hand (Figure 1.16(a)). f) *Palmprint capture*. The user is asked to put his/her hand in an encasing having controlled illumination and a platform that restricts hand motion (Figure 1.16(c)). An image is then captured using a scanner or a camera. g) *Gait Recognition*. The person is recognized based on a video recording the walking patterns. h) *Other biometric Modalities*. Other biometric modalities such as ear and odor are difficult to capture and need a dedicated setup.

The dedicated image capturing devices suffer from the following shortcomings:

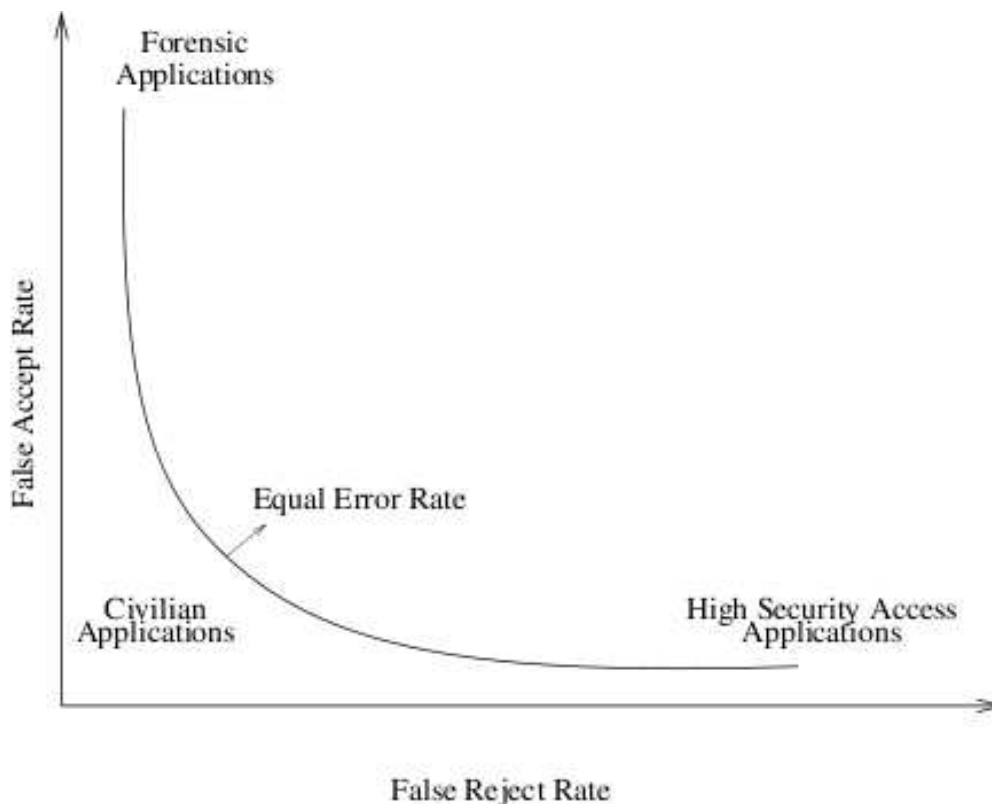


Figure 1.17: Effect of EER on user acceptance of various Biometric System.

- High maintenance costs,
- It is not intuitive to use and hence cannot be used in unsupervised scenarios,
- Once the system is installed, it is difficult to add and integrate more biometric modalities for use in a multi-biometric system,
- It assumes a co-operative user,
- Hygiene based concerns are now being raised for some of the traditional contact based biometric capture devices(like fingerprint, palmprint etc.),
- These devices are more expensive compared to the new cameras that are becoming cheaper with the day and can be explored as an alternative image capturing device.

The state-of-art Biometric Recognition methods had focused on studying and developing techniques for better storing of templates upon enrollment and improvement of recognition accuracy for better identification (Refer to Section 1.5). For all practical purposes, a standard robust Biometric Capture device is used that can generate images in ideal laboratory like conditions. The Biometric Capture module was left untouched for a long time until its revival recently. The reason being an exponential downward surge in the costs of camera devices, because of which high resolution cameras can be obtained very easily now. Moreover, people are becoming increasingly aware of the potential advantages of using camera based imaging which are the following:

- It is intuitive to use and hence can be used in an unsupervised manner,
- When used as a touch-less imaging system, it has high user acceptance as it is free from hygiene concerns,
- Another major advantage is its potential for use in multi-biometric systems, as a single camera device can be used to capture more than one biometric (say, face and palm, or palm and fingerprints etc.),
- It can be used in surveillance videos and scenarios when the subject is not co-operative, and
- It is also easy to add and integrate more biometric modalities into an already functioning system with minimum overhead as and when it is required.
- As already mentioned, it is inexpensive.

Hence, the easy availability of camera devices now-a-days combined with its potential to overcome all the disadvantages of traditional image capture devices has made camera based imaging a lucrative alternative to sensors and scanners. The growing trend of using camera based devices can be classified into broad categories. **a) Under controlled or restricted surroundings** and **b) Under uncontrolled surrounding conditions**. The first category basically just replaces a sensor and/or scanner in its role as an image capturing device. The rest of the controlled setup and the restrictions imposed on the user need not change. The advantages in such a case may not constitute the entire set described above. The major advantage in such a case is the reduced initial installation cost, more control over the resolution settings in a dynamic manner, and getting rid of hygiene

based concerns when used as touch-less imaging system. This is mostly employed in applications that need stringent security measures where user comfort is not of prime concern for eg. forensic applications, airport security checks, controlling access to documents and bank accounts of high importance etc. The other method describes a generic scenario when a user just presents the biometric in front of a camera that captures it. The setting is highly intuitive for use and has high user acceptance. This is generally encountered in commercial civilian application (unlocking module in laptops and mobile devices), biometrics based attendance systems (in classrooms and medium sized organizations), surveillance videos etc.

Many researchers are now focusing on developing new recognition algorithms using generic camera based imaging. The most researched upon area in this direction is probably Automatic Face Recognition (AFR) because of obvious ease of photographing face using camera ([18]). There has been some work to develop view invariant gait recognition algorithm as well ([28]). In this thesis, we develop similar solutions for palmprint recognition. Previous research reported on unconstrained palmprint imaging includes solutions for imaging without a platform (or touch-less imaging) ([36, 37, 38]). But the distance between the hand and the camera is required to be fixed. Julien *et al.* proposed a new algorithm, which did not require the distance between the palm and the camera to be fixed but needed users to show their palms parallel to the imaging plane of the camera. This is the first time that the problem of developing solutions for truly unconstrained and unsupervised palmprint imaging having varying poses and differing illumination is being attempted. The exact problem statement has been given in the following section.

1.9 Problem Statement

As noted in 1.1, due to the ever reducing camera sizes and equipment cost, vendors are now actively integrating camera devices in laptops and mobile devices. This has opened up new vistas for the applicability of biometrics based recognition devices in civilian applications. This is a great opportunity to bring the biometric technology to the common man. However, there are many challenges to be addressed to bring the biometric softwares up-to-date with this new trend. Even though the cameras available now-a-days capture images that are aesthetically appealing to the human eye, the information captured is not sufficient from the point of view of a biometric recognition application. Even small degradations cause the recognition accuracy to go down considerably for almost all biometric modalities. Hence, from the usability point-of-view of a recognition algorithm, the cameras capture noisy low-quality images only. Another issue is that of capturing unsupervised biometric data in uncontrolled situations leading to pose and illumination variations. The onus then lies on the creation of intelligent algorithms to extract useful information for person identification.

In this thesis, we develop a biometric recognition algorithm with a generic web camera used as an image capturing device and unsupervised and uncontrolled surrounding conditions. The biometric modality that we choose to work with is palmprint (See Section 1.9.1). We aim to systematically quantify and study the effects of the following factors affecting the recognition accuracy of the palmprint recognition system.

- **Variations caused due to unsupervised user behavior during image capture:** Pose variations due to leniency on the part of user at the time of presenting biometric degrade the image quality resulting in an increase of the intra-class distances.

- **The use of a low-end camera in uncontrolled surrounding conditions:** cause problems of illumination variation, jitter due to hand motion, noise etc. Hand motion and poor camera quality result in out-of-focus images whereas illumination variations and shadow creation result in false line impressions which need to be filtered out effectively.

We aim to study the effect of these variations on the system and also provide solutions to the above mentioned problems to create a robust camera based palmprint recognition system. The applications of such a system include access control devices at entry points in medium sized organizations and unlocking modules in laptops and mobile devices. This can also be used as an add on for creating multi biometric systems.

1.9.1 Palmprint as a biometric modality

Palmprint is the inner part of a person's hand. For centuries, the palm line patterns have popularly been believed to be able to predict a person's future. But its uniqueness and capacity for distinguishing individuals has come to fore only recently. In [39], Nicole Duta *et al* showed this for the first time. It is a physiological biometric considered to be a weak biometric(weak biometric means that the accuracy is approx. 99% as compared to a strong biometric having accuracy in the range of 99.99%). Palm Lines, which are the most distinguishing features of a palm, are coarse and hence very easy to capture using a low resolution camera. Palmprint rates high on user acceptability since it is not used for applications enforcing stringent security measures and also because the hand is the most mobile organ of the body rendering the person more in control, and hence it is easy to present in a touchless imaging system. Hence resolving the hygiene related concerns. There is an added advantage when using palmprint as a biometric. It is very easy and convenient to integrate palmprint into an already existing Biometric Recognition System since it does not need a dedicated capture device. Because of the presence of coarse distinguishing lines, it is possible to capture palm lines even at a low resolution, using a digital camera as shown in [8]. This feature makes it very useful for use in a camera based modality as cameras capture only low resolution details when compared to scanners. Moreover, due to easy mobility of hand, it is very easy to pose palm. All this combined with a moderate recognition accuracy on large datasets makes palmprint the ideal choice as an add-on in a multi biometric system.

All the factors defined above make palmprint a very useful biometric for camera based biometric recognition devices. Although, there has been steady growth in the direction of developing a contactless palmprint recognizer(Section 2.1.1), a truly pose invariant and robust palmprint recognition system is yet to be proposed. This makes pose and illumination invariant palmprint recognition a very interesting study.

1.10 Contributions of the thesis

In this thesis, we have proposed a robust and practical design for an unconstrained palmprint recognition system. The system proposed by us deals with real life data, modeling accurately the variations in illumination, pose and scale. We provide solutions to two major problems faced while using a generic camera for uncontrolled camera devices in an unsupervised scenario. The main contributions of this thesis can be listed as follows:

- **Achieving pose invariance in palmprint recognition.**
- **Achieving illumination invariance in palmprint recognition.**
- **Palmprint Dataset:** Presenting two new palmprint datasets having variations in pose and illumination.

For data capture, we used a generic day-to-day camera like a web camera which can be easily found with everyone these days, and has no special focusing ability or exceptionally high image resolution. We choose to work with this kind of a camera because it is easily available mounted on laptops or camera phones and this is the kind of camera that an average person will use to authenticate his/her identity while using a laptop or a mobile camera. These cameras generally produce low resolution and noisy/grainy images which are very difficult to work with especially in a recognition based algorithm. The problems are compounded by the fact that there is no supervision of users, when they present biometric modalities in a generic fashion to these cameras. There is no way to fix/control lighting conditions, shadows cast by surrounding objects, background conditions etc. Hence, to make biometric recognition truly accessible for practical use to a common man today, we must develop algorithms that develop smart algorithms, since hardware alone cannot provide the solution here. This thesis is a step forward in that direction.

1.11 Thesis Overview

We have presented an overview of the problems, techniques, modalities and applications of biometric authentication. We have also presented the specific problems that we are addressing in this thesis.

In Chapter 2, we talk about the specifics of the design and algorithms used in the state of the art palmprint recognition systems. A literature survey of the various techniques in use has been provided with discussions after which we would like to draw your attention to the novelty and relevance of our problem domain in the chosen field with reference to the ongoing research. A set of challenges faced is then presented followed by a discussion on the popular solutions in practice and their applicability to our problem. The chapter concludes by highlighting the specific problem areas for which we need to develop algorithms. Chapter 3 and Chapter 4 then describe the proposed solution in detail. We present the algorithm followed by experiments and results. Chapter 5 presents concluding remarks and future work.

Chapter 2

Literature Survey

In the last chapter, we presented an overview of biometrics. Here, we discuss about the palmprint recognition process in detail. We start with a survey of the state-of-art methods of palmprint recognition. The section provides a step by step procedural description of the recognition process and the various techniques and methods prevalent to achieve required results in literature. Then, we analyze the applicability of these steps in the unconstrained camera based imaging. Since the image capturing process differs, not all the steps are directly applicable. We identify the challenges faced by us in our problem area and try to find parallels and possible solutions by surveying the research done for similar problems in biometrics and other domains. We explain the state-of-art methodologies, and why most of them are not directly applicable to our problem. The challenges and bottlenecks of this problem and the related research work and reasons as to why they cannot be used shall also be made clear in this chapter.

2.1 Palmprint Recognition

Palmprint:

In this thesis, we refer to that area of palm, also known as “lower palm” by some, having the principal or major lines([40]). Human palm has line features including minutiae points just like in the case of fingerprints. Like fingerprints, a palm can be scanned to obtain an abundance of ridges and minutiae which form the finer details of the palm([41]). Some vendor systems using these finer details for commercial use have been reported ([42],[43]). These finer details are obtained by scanning the palm image at a high resolution of 1000 DPI ([40]). In this thesis, we focus on the higher level textural information present on the palm in the form of major lines and small wrinkles. The reason for choosing to work with these is that the higher level details can be captured using a generic web camera at low resolution. The loss of accuracy on account of losing out on finer details is compensated by the gain in user convenience and no extra requirements of hardware. The accuracies obtained by using just the high level information is found to be sufficient for recognition of small groups of individuals (as shown by the literature survey in this section).

The literature survey, hence, pertains to work done only on palmprint recognition using higher level texture details. Any palmprint recognition process has the following major steps:

- Image Capture

- Image Preprocessing,
- Feature extraction,
- Matching.

The functionality of these has been described below:

2.1.1 Image Capture

It is a very important step in the biometric template recognition as the quality of image captured has a great impact on the recognition accuracy. Initially, the palm images were obtained after scanning an inked palm [39, 44]. An inked image was able to capture all the major palm lines as well as the minor creases and wrinkles. Scanning it resulted in a high resolution image having distinctive signatures unique to people. Inked palm images can be used for forensic applications easily. However, the technique is not a good way of image capture for commercial applications. This led to creating dedicated scanners for palmprints. These scanners need to be bigger than the traditional fingerprint scanners in accordance with the size of the palm. This was convenient for use in offline-palmprint recognition[45, 46]. However, scanning a palm image takes a few seconds(because of its size) which makes it difficult to be used with an Online palmprint recognition system in real time.

It is possible to capture the principal lines and a few major wrinkles in a palm image even at low resolution. One of the advantages of using palmprint as a biometric is that it can be used to distinguish people even with this low resolution image. This enables Online palmprint based identification in real time applications. David Zhang *et al.* [8], introduced camera based palmprint recognition using CCD cameras based on a special device for online palmprint acquisition. A stable palmprint image is obtained by using a case and a cover that are used to form a semi-closed environment. A ring source provides uniform lighting conditions during palmprint image capturing. Also, six pegs on the platform serve as control points for the placement of the user's hands [47].

2.1.2 Image Preprocessing

The function of this stage is to segment out the hand image from the background and obtaining a binary image of the hand with known orientation and geometry. It consists of the following steps:

- Extracting hand from the background and Image Binarization,
- Determining the orientation of the hand, and
- Detecting key points.
- Establishing a co-ordinate system
- Extraction of Region Of Interest(ROI) i.e. palm.

Extracting hand from the background and Image Binarization

. This step is similar in almost all preprocessing algorithms. The images obtained from the setup described in 2.1.1 have the following properties:

- The background is fixed.
- The distance between camera and the palm is fixed.
- Since the palm is placed on a platform having fixed pegs, there are no pose variations.
- Due to the artificial lighting present in the enclosing setup, the illumination is uniform and remains constant over sessions.

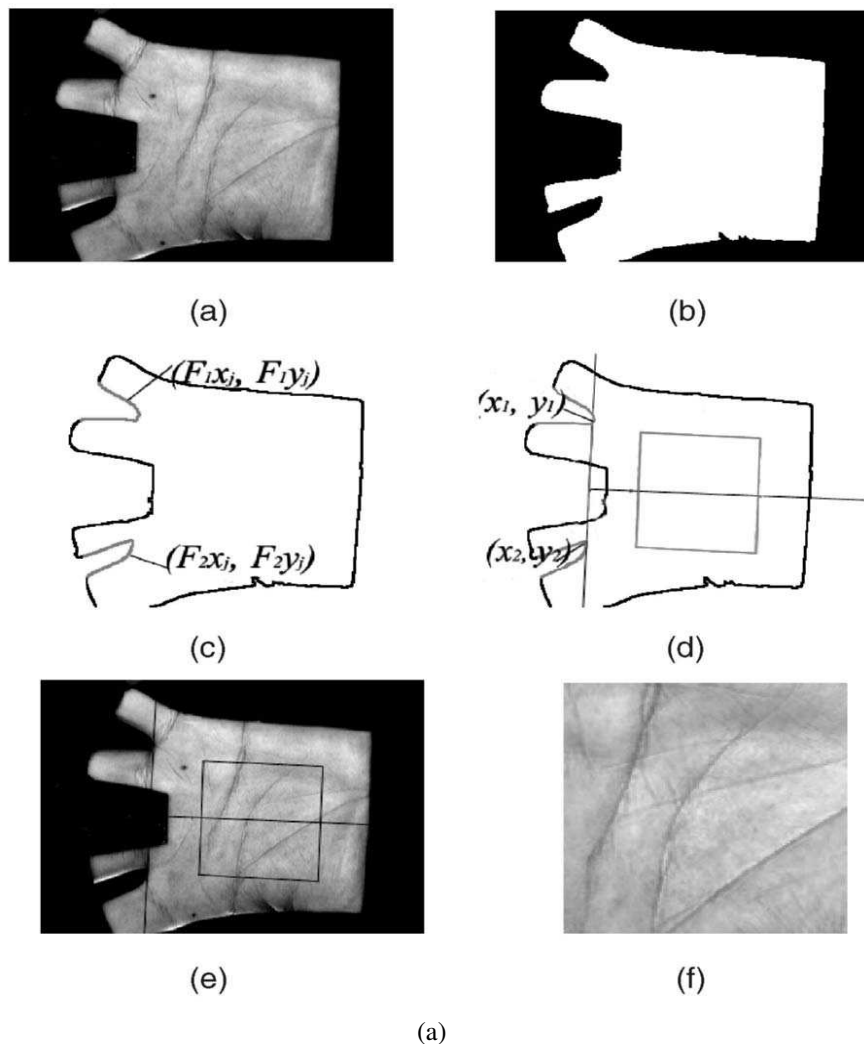


Figure 2.1: The main steps of preprocessing and palm extraction. a) Original Image, b) Binarized image, c) Boundary Tracking, d) Fixing the co-ordinate system, e) extracting the central region as ROI from the image, and f) resulting extracted palm

Due to the fixed background and the uniform illumination with previously known parameters, it is easy to segment out the hand using Otsu's method [8]. Otsu's method chooses the threshold to minimize the intraclass variance of the two groups of pixels (the constant background and the hand image) separated by the thresholding operator [48]. The segmented out hand is then represented with white pixels for the purpose of computation of finger valleys.

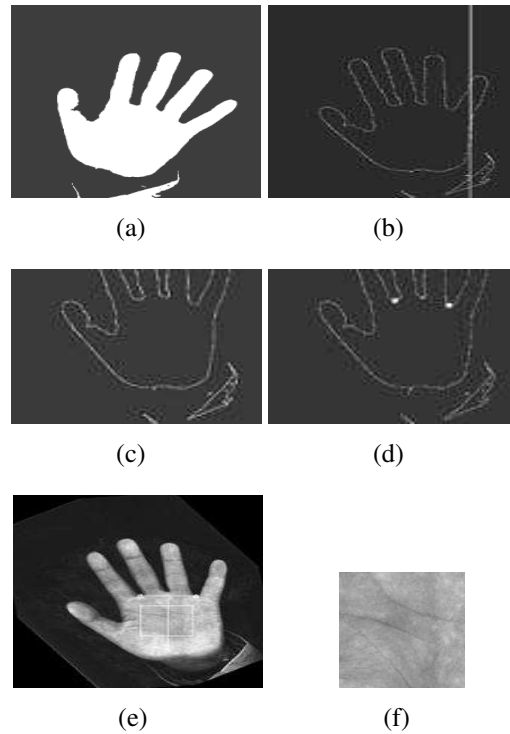


Figure 2.2: The main steps of the modified preprocessing method. (a)binary image, (b)Rotation basing on outer boundary direction and obtain the edge image. (c)truncate the boundary image with 8 intersections (d) trace finger gaps (e) re-rotation of the original image based on the coordinate based on gaps between fingers (f) extract a 128 128 sub-image.

Determining orientation of the hand

Knowing the orientation of the hand is important for detecting the location of key points on the hand. There are several ways to find out the orientation of the hand. One way is to use pegs attached to the platform during image capture. These pegs can easily be identified in the image and used as landmarks [47, 8] which help in fixing the orientation of the hand. Some methods explicitly rotate the image to fix the rotation. The image is rotated such as to align the horizontal axis with the line joining the finger valleys formed by the index finger and the middle finger and another formed by the middle finger and the ring finger [49] (see Figure 2.2). Another popular approach approximates the shape of the hand to an ellipse. The orientation of the hand is then approximated by the major axis of the best fitting ellipse [34].

After the orientation of the hand is determined, an orthogonal co-ordinate system is setup by finding the keypoints/finger valleys between fingers (See Section 2.1.2). There are three ways to

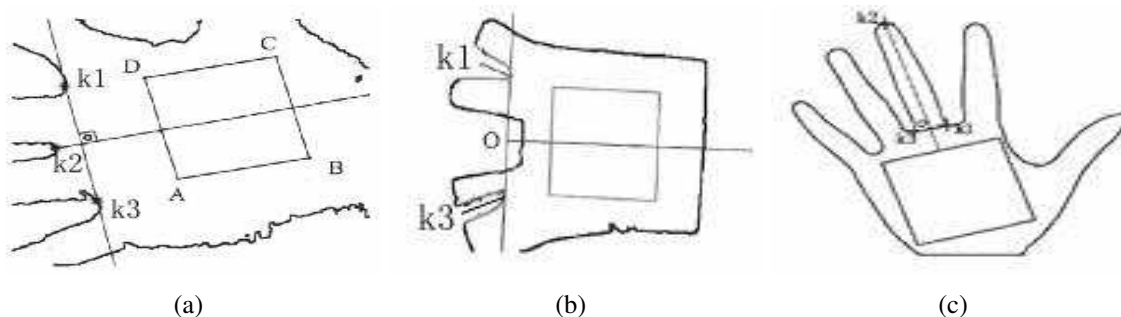


Figure 2.3: The three approaches for fixing co-ordinate axes for Palm Extraction.

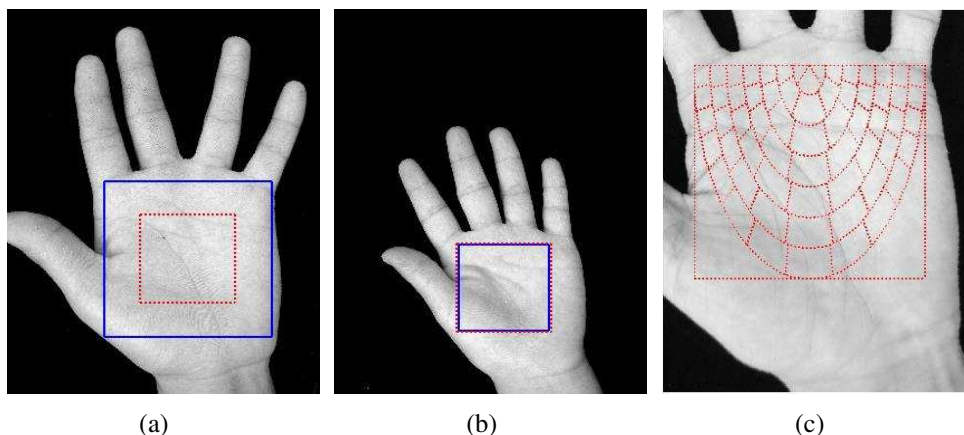


Figure 2.4: Figure a) and b) show the difference between various palm sizes. Figure c) shows the elliptical palm rings extracted from the hand.

fix the co-ordinate system[50]. The first approach defines three keypoints k_1 , k_2 and k_3 , as shown in Figure 2.4. The Y-axis of the palmprint co-ordinate system is the line joining k_1 and k_3 . Then, X-axis is the line perpendicular to Y-axis and passing through k_2 , and their intersection point is the origin ([51]). In the second method, the approach is similar to the first one but the points k_1 , k_2 and k_3 are defined differently as shown in Figure 2.4. In the third approach, only two points are used to determine the co-ordinate system. Y-axis is the line joining k_1 and k_3 ([52]). X-axis is taken to be the line perpendicular to Y-axis and passing through the mid-point of the line joining keypoints k_1 and k_3 ([8]). (See Figure 2.4)

Detecting keypoints between fingers

The finger valleys (depression points present between fingers) are chosen to fix a co-ordinate system with respect to which the palm image can be extracted. The line joining the finger valleys is chosen to be the Y-axis and the X-axis is the line perpendicular to it. Hence, it is very important to find out the finger valley points correctly to reliably match two given palm images. There are different ways for implementing finger valley detection. These can be classified under: a) Tangent-based, b) Bisector-based, and c) Keypoint finding. They have been described below.

Tangent Based and Bisector Based: As described in [8], the process of finger valley detection involves the following steps:

Step 1. Apply a lowpass filter, $L(u, v)$, such as Gaussian smoothing, to the original image, $O(x, y)$. A threshold, T_p , is used to convert the convolved image to a binary image, $B(x, y)$ (See Figure 2.1b).

Step 2. Obtain the boundaries of the gaps, $(F_i x_j, F_i y_j)$ ($i = 1, 2$), between the fingers using a boundary tracking algorithm as shown in Figure 2.1c (see details below). The boundary of the gap between the ring and middle fingers is not extracted since it is not useful for the following processing.

Step 3. Compute the tangent of the two gaps. Let (x_1, y_1) and (x_2, y_2) be any points on $(F_1 x_j, F_1 y_j)$ and $(F_2 x_j, F_2 y_j)$, respectively. If the line $(y = mx + c)$ passing through these two points satisfies the inequality, $F_i y_j \leq F_i x_j + c$, for all i and j (See Figure 2.1d), then the line $(y = mx + c)$ is considered to be the tangent of the two gaps.

Step 4. Line up (x_1, y_1) and (x_2, y_2) to get the Y-axis of the palmprint coordinate system, and use a line passing through the midpoint of these two points, which is perpendicular to the Y-axis, to determine the origin of the coordinate system (See Figure 2.1d).

Step 5. Extract a sub-image of a fixed size based on the coordinate system. The sub-image is located at a certain area of the palmprint image for feature extraction (See Figure 2.1e and Figure 2.1f). This distance is fixed across images for better accuracy.

Bisector Based: Bisector-based approach constructs a line using two points, the center of gravity of a finger boundary and the midpoint of its start and end points. The intersection of the line and the finger boundary is considered a key point.

Keypoint Detection: In this approach, the axis formed is the line joining the two key points instead of the explicit tangent approach. Much work has been reported on the methods used to detect the finger valleys from the contour (boundary) of the hand. This is done primarily for use in hand geometry based recognition [53]. In hand geometry based recognition, geometrical landmarks like the fingertip points, the valley points between fingers, length and width of fingers etc constitute the feature vector. These landmarks are extracted by travelling along the hand boundary and searching for curvature extremities and corner point detection [53].

‘Corner’ or ‘dominant point’ detection is important for pattern or picture analysis. A corner is (informally) defined as a high-curvature point on a simple digital arc or curve. Corners can be used to segment arcs or curves. We consider curves ρ in the digital plane. Pixels p_i in such a digital curve $\rho = p_0, p_1, \dots, p_{n-1}$ have coordinates (x_i, y_i) . In order to detect a corner at pixel p_i on a curve ρ , it is common practice that a corner detector considers an angular measure based on a predecessor p_{i-b} , p_i itself, and a successor p_{i+f} , where $b, f > 0$ are fixed or variables within a defined interval.

Corner detection is important task in hand image analysis. Corners are important image hand features since they correspond to unique features of and hand, and are invariant to many transformations. A corner point on a given hand contour sub-tends a sharp angle between its neighboring points. A discrete boundary curve is described by $r_k = [x_k, y_k]^t$, $k = 0, 1, \dots, n - 1$ where t is the matrix response, and x_k and y_k denote the x and y coordinates in the $2D$ image plane.

In first step, the curve is smoothed to prevent variations in the estimated angular values:

$$\tilde{x}_k = \frac{1}{2w+1} \sum_{l=k-w}^{k+w} x_l, \quad \tilde{y}_k = \frac{1}{2w+1} \sum_{l=k-w}^{k+w} y_l \quad (2.1)$$

where $\tilde{(\cdot)}$ is smoothed value of (\cdot) and w is a small integer number.

The angle Ψ_k associated with each curve point r_k is then computed as the angle between the two vectors $(\tilde{r}_{k+1}, \tilde{r}_k)$ and $(\tilde{r}_k, \tilde{r}_{k-1})$ using the smoothed curve points:

$$\tilde{\Psi}_k = \cos^{-1}\left(\frac{a^2 + b^2 - c^2}{2ab}\right) \quad (2.2)$$

where $a = |\tilde{r}_{k-1} - \tilde{r}_k|$, $b = |\tilde{r}_{k+1} - \tilde{r}_k|$ and $c = |\tilde{r}_{k+1} - \tilde{r}_{k-1}|$.

Significant corners are extracted from those points which have the local minimum angular values.

Palm Extraction

Once the co-ordinate axes are fixed, extracting the palm is relatively straight forward. The most common approach is to extract a square region of fixed size from a palm. As shown in Figure 2.1e), one of the way to extract a palm is to extract a square sub-image of fixed size taken at a fixed distance from the axis [8]. But this can be a problem if the size of user palms vary greatly. If the fixed size is high, it might be too big for a small hand (for eg. a child's hand); if it is small, on the other hand, it may not be able to extract complete information from a user with a bigger hand (see Figure 2.4(a) and 2.4(b)). Poon *et al.* ([54]) and Wang *et al.* ([50]) proposed capturing a variable sized palm so as to capture the maximum information from a user's hand. To extract features, Wang *et al.* proposed normalization of the palm image to a fixed size (scaling down). That way all the information in the palm could be imbibed even though on a smaller scale. Poon *et al.* suggested extracting a fixed length feature vector by dividing the variable sized palm into half elliptical rings of various radii as shown in the Figure 2.4(c). Each ring is then divided into equal number of regions and hence, getting a feature vector of uniform length.

2.1.3 Feature Extraction and Matching

Once the central part of the palm is segmented out, features can be extracted for matching. Verification based algorithms can be classified as line-, subspace- and statistic-based. A detailed study of these has been provided in [55]. These algorithms have been described below.

Line Based Approaches

Line based approaches either develop edge detectors or use existing edge detection methods to extract palm lines. Wu *et al.* use Canny edge operator to detect palm lines. The orientations of the edge points are passed into four membership functions representing four directions. For each direction, $E_{R,i} = \sum_{(x,y) \in R} (Mag(x,y) \times \mu_i(x,y))^2$, where μ_i represents one of the membership functions; Mag represents the magnitude of the lines and R is a local region. The feature value, $E_{R,i}$ is normalized. Finally, Euclidean distance is used for matching.

Pedro *et al.* employ Sobel masks to enhance edge information and the statistical information in the processed images is used to estimate an optimal threshold for extracting the edges [56]. The authors then utilize a thinning algorithm to further process the edges. Several descriptors of the edges are computed as features for matching.

Kung *et al.* proposed a feature vector based on a low-resolution edge map. The feature vector is passed into decision-based neural networks. This was the first paper to report an on-line palmprint recognition method [57].

Huang *et al.* presented a two-level modified finite radon transform and a dynamic threshold to extract major wrinkles and principal lines. Two binary edge maps are compared based on a matching scheme called pixel-to-area comparison [58]. Even though some strong wrinkles are included in the edge maps, the major features in this method are principal lines, which are genetically dependent [59].

Subspace Based Approaches

Subspace-based approaches also called appearance-based approaches in the literature of face recognition. This is the name given to the group of algorithms that use Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) and Independent Component Analysis (ICA) ([60, 61, 62]). The features are the subspace coefficients. Various distance measures and classifiers are used to compare the features. Researchers have also applied wavelets, Gabor, discrete cosine transformation (DCT), and kernels in their methods ([62, 63, 64, 65]). In general, subspace based methods do not make use of the prior knowledge available about palmprints.

Statistical Approaches

These are either local or global in nature. Local statistical approaches divide the image into smaller sub-images in the transformed domain (transformations are usually done to transform the image from time domain to frequency domain or extracting some other features like DCT coefficients etc), and then obtain features on each of these sub-images to create a feature vector. Statistics like mean, variance etc of each of these smaller regions forms the feature vector here ([66, 54]). The applied transforms can be Gabor, wavelet etc. Wang *et al.* have proposed the use of histograms of local binary pattern as features. On the other hand, global approaches compute statistical features at a global level directly from the transformed images ([67]) instead of dividing the image into smaller sub parts. Moments, center of gravity and density have been used as global statistical features.

Other Approaches

Some approaches are difficult to classify because they employ a combination of image processing techniques and a variety of properties to obtain the final feature vector.

Doi *et al.* regard the intersection points of finger skeletal lines and finger creases and the intersection points of the extended finger skeletal lines and principal lines as feature points [66]. Han extracted seven specified line profiles from pre-processed image and combined it with features from the fingers and used wavelets to compute low frequency information. The dimensionality of this new feature vector is reduced using PCA. Finally, generalized learning vector quantization and optimal positive Boolean function are used to make final decision.

Koichi *et al.* also propose a correlation approach [68]. The amplitude spectrum of two segmented images is used to estimate their rotational and scale differences. One of the images is rotated and scaled and then their amplitude information in the frequency domain is removed.

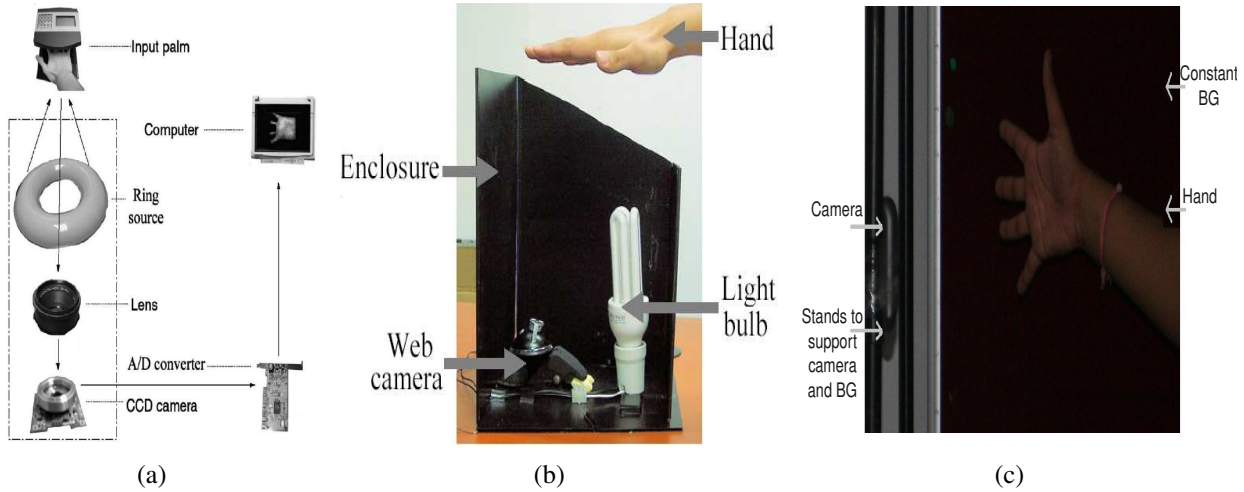


Figure 2.5: Different image capturing setups. a) Constrained image capture, b) Touchless Imaging, c) Proposed unconstrained touchless capture

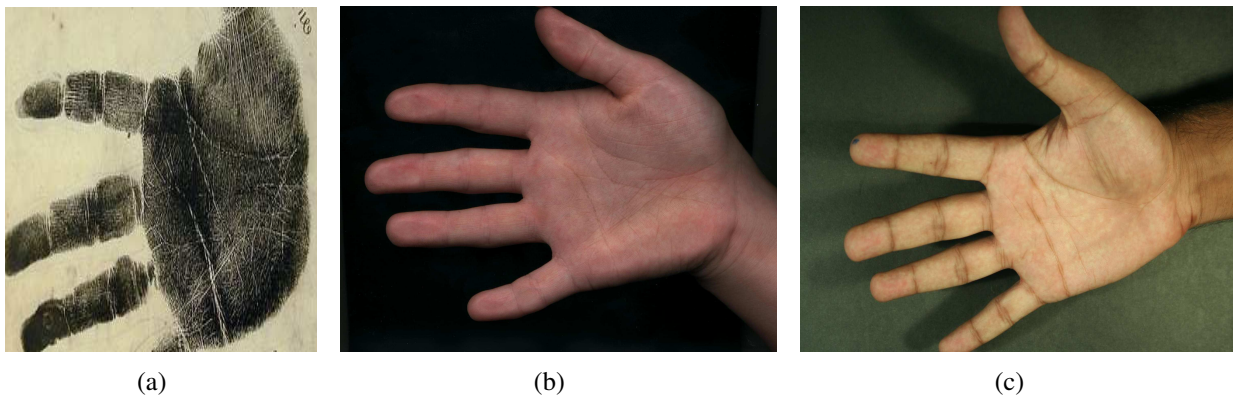


Figure 2.6: Different palm images obtained by using different setups. a) Inked palm, b) scanned palm, c) palm image obtained by unconstrained imaging.

Finally, band-limited phase-only correlation (BLPOC) is used to compute the similarity of two images. BLPOC only considers low to middle frequency information.

Zhou *et al.* [69] employ wavelet to decompose palmprints and use support vector machine (SVM) as a classifier. The input of the SVM is low sub-band images. This approach may overlook some important information in the middle frequency spectrum.

2.2 Towards Unconstrained Camera Based Palmprint Imaging

In the previous section, we had a look at the literature survey of the state-of-art palmprint recognition techniques in use currently. One can see that the image capturing setup used in the methods described above in 2.1.1 is not suitable for use with laptops and mobile devices. The bulky setup is difficult to carry around and would be a hindrance. Hence, we need a light capturing device which can be easily mobilized along with these systems. The digital web cameras attached to a

laptop and mobile are in an ideal position for capturing the required templates. This has already been used for facial image capture for face recognition based access in Lenovo laptops [70]. We could use the same camera for palmprint imaging. Palmprint would act as a suitable add-on for a multi biometric setup already using some other modality(for eg. face) since it would use the same capturing device. It could also act as a stand alone application.

As already mentioned in Chapter 1, unconstrained imaging has a set of associated challenges viz. pose transformations, illumination variations etc. Once we have a design and the hardware in place, we need to have reliable algorithms to make palmprint recognition with unconstrained camera imaging actually possible. Previous research reported on unconstrained palmprint imaging includes solutions for imaging without a platform (or touchless imaging) ([36, 37, 38]). But the distance between the hand and the camera is required to be fixed. Julien *et al.* proposed a new algorithm, which did not require the distance between the palm and the camera to be fixed but needed users to show their palms parallel to the imaging plane of the camera. This is the first time that the problem of developing solutions for truly unconstrained and unsupervised palmprint imaging having varying poses and differing illumination is being attempted.

There have been attempts for creating similar systems having unconstrained imaging for other biometric modalities like face([18]) and gait ([28]) etc. But the design for each modality is different and unique. It depends on the characteristics of the biometric modality under consideration. The solution for illumination invariant face recognition uses manifolds to represent faces. View invariant gait recognition uses the 3D structure of the human body to determine information about the view. Both these features are absent in palmprint, which has very low textural information relatively and hence cannot be transformed in a new feature space. Its 3D structure also does not reveal much information. Hence, we need to look at palmprint recognition in unconstrained imaging scenario as an independent problem. The challenges faced in this problem have been described in the next section. We then present the solutions to similar problems as solved in related problem areas and analyze the aspects related to our problem. We effectively survey the literature and bring out the need for novelty in our solution and present our solution thereafter.

Please notice the different image capturing setups in Figure 2.5. The first image shows the state-of-the-art constrained imaging capture where the subject is asked to place his hand on a fixed plate. The second image shows a touchless imaging capture setup where the palm is required to be kept parallel to the camera imaging place. The third figure shows the setup proposed by us. The user's palm can be presented in a truly unconstrained manner. The different palm images captured using different systems is shown in Figure 2.6. The first image shows the inked palm. The second image shows the images captured by a scanner. The third image is the image captured using an unconstrained palm capture device.

2.3 Unconstrained Palm Imaging: Challenges and existing solutions

State-of-art palmprint recognition algorithm works with images captured in an ideal semi-closed environment. Thus, the solution used for those images varies greatly than the solution needed for images captured using unconstrained and unsupervised cameras. To come up with an effective solution for the problem at hand, we must first list out the challenges faced and identifying the



Figure 2.7: Challenging cases of unconstrained camera imaging



Figure 2.8: Segmenting hand image from a cluttered background.

bottlenecks of the existing solution in tackling them. Hence, the various challenges faced in this technique and details about the existing solutions are presented below.

2.3.1 Background Clutter

In an unconstrained image capturing setup, the idea is that the constraints imposed on the subject are as little as possible. For this, we use a fixed camera, and the subject is asked to pose a stretched palm intuitively. In an unconstrained image capturing setup, the idea is that the constraints imposed on the subject are as little as possible. For this, we use a fixed camera, and the subject is asked to pose a stretched palm intuitively. In an unconstrained image capturing setup, the idea is that the constraints imposed on the subject are as little as possible. For this, we use a fixed camera, and the subject is asked to pose a stretched palm intuitively in front of the camera. Very few subjects tended to place their hands parallel to the camera plane unless explicitly asked to do so. In the most generic case, there are no restrictions on the background either. Although it is not a necessary requirement for our purpose, we describe it here for the sake of completeness. Various algorithms have been reported in literature to detect and track the location of hand image in videos containing hand images against a variety of background clutter [36, 71]. Some algorithms track hand placed against a cluttered background using single images only [72]. The use of skin and hand shape filters is used for this. A detailed survey of methods of hand segmentation for the use of gesture estimation for Human-Computer Interaction (HCI) has been provided in [73]. Segmenting hand placed against a background is a sub problem of interpreting hand gestures. Some images showing segmented hand images taken from [72], and [71] are shown in Figure 2.8.

Thus, the problem of detecting hand placed against a cluttered background in real-time is solved to a large extent, and can be, as such, plugged directly into a palmprint recognition algorithm

as a preprocessing step. Since the scope of this work concerns providing a solution to the pose invariance and illumination problem, we used a constant background which allows us to extract the hand image from the background in a straightforward manner. Our setup has been shown in Figure 2.5(c). A couple of different poses to be matched are shown in Figure 3.1. As can be seen, the background is chosen to be constant here.

2.3.2 Feature Extraction

All weak and strong palm lines, form the set of features used for palm comparison. A lot of research has already gone into finding out the best set of palm features to be used. These approaches can broadly be categorized into these three categories: a) texture based b) line based and c) appearance based, which have already been described in Section 2.1.3. However, these approaches have been described considering the images with restricted pose. Since, most of these features are not projective invariant (eg. eigenvalues and line-based methods), they are not suitable to be used for images with projective pose transformations. Hence, we need to analyze what features will be best suited for our purpose.

2.3.3 Scale Variations

One of the variables encountered during unconstrained camera imaging is the distance between the camera and the subject's hand. Earlier, the fixed plate also helped in keeping this distance constant, so the obtained image is always the same size. One must also address the problem of scale variation when making a touchless palmprint recognition system ([36, 37, 38, 74]). But as mentioned earlier, subjects are asked to keep their hands parallel to the camera imaging plane in the above work(s). In that case, the scale in both X- and Y- dimensions is the same, and the problem reduces to finding a constant scale factor such that any two images when being compared, are of the same size. This can be done by either scaling all images to a standard pre-fixed size ([36, 37]), or to scale the images in ratio to some hand geometry measurement eg. palm width ([74]). But in the case of images having pose variations, scaling factor is not a constant value. Hence, it is not possible to directly solve for scaling. It becomes a part of solving for pose transformation variable.

2.3.4 Pose Variation in images

Solving pose variation for palmprint has never been attempted before. This is especially difficult to solve for in the case of palmprint since palmprint is a comparatively low textured image. The information available is only in the form of palm lines and wrinkles. The characteristic properties of lines and their layout on the palm changes extensively when subjected to projective transformation. As a result of which, it is difficult to find landmark points crucial for pose estimation and creating projective invariant features. This is essentially a problem to find the best possible way to find landmark points with maximum reliability. This becomes trickier because of the non-uniform lighting on hand as explained below.

2.3.5 Illumination Variations and Noisy Artifacts

ILLUMINATION VARIATION: Comparing two different palm images captured in different illumination settings is a difficult thing to do. Since, it affects the line extraction algorithm. Especially, in the case of natural lighting, there can be a lot of difference between an image captured at day time and one captured at night. This disparity can be controlled to an extent by using an artificial source of light attached to the camera. Since we are dealing with pose variations here, the light reflected from the hand is affected by the following factors: a) the light coming from the other sources present nearby, and b) specularity of the skin. These create an impression of non-uniform lighting on the hand, because of the different depths of various points due to the variation in pose. The visibility of some lines is greatly affected on changing hand pose as can be seen in Figure 3.1 and Figure 2.7.

Thus, when talking about pose variation, even with constant/artificial lighting, we must handle cases of illumination variation to find a holistic solution. It must be noted here that even the light source attached to the camera is used without the encompassing box used to control the effect of other light sources in the vicinity, which end up increasing the volume of the overall image capturing device([37]).

NOISY ARTIFACTS:When hand is turned at different angles, stretching of skin results in new wrinkles surfacing up. This can become a real problem when building a pose invariant recognition system, since there is a different set of wrinkles associated with different poses. Thus, when posing the hand intuitively in front of the camera, extra wrinkles are observed to be introduced because of the following reasons: a) Different people hold out their thumbs differently at different times. This causes wrinkles of different intensity to be introduced at the thumb and palm joint. Hence, the intensity of the line found at this position may differ. b) While the hand is tilted, it may not always be possible that all the fingers lie in the same plane, depending on how loosely/tightly the subjects hold out their hand. This can also create false line-like features. c) Due to the tilting of the hand, it is seen that shadow of the hand boundary may fall on the hand itself. This leads to an illusion of a line being present at that position on account of intensity difference(See Figure 2.7(c),2.7(d)).

In all the above mentioned scenarios, either we end up detecting lines that are not there or else missing out the weaker lines that are supposed to be recognized because of the improper lighting. Apart from line detection, improper lighting condition also interferes with hand image binarization, because of the variations in skin reflectance[75]. We propose a new method to approach the solution to both these problems in this article.

2.4 Unconstrained Palmprint Imaging: Bottlenecks

So far, we have seen all the challenges faced with unconstrained palmprint imaging and how the existing solutions can be used to solve some of them. A brief summary is presented here. Extracting hand from a cluttered background is a solved problem (Section 2.3.1). A lot of work has been done on various methods of feature extraction. Gabor filter results in pretty robust features (Section 2.3.2). Some methods to solve scale variations have also been reported. But their scope is limited to naive scaling of hand, which works well only when the palm is almost parallel to the imaging plane (Section 2.3.3). Scale variations when palm is not parallel to the camera imaging plane can only be truly handled using pose transformation parameters. The problem of pose, scale and illu-

mination invariance has never been attempted in literature so far to the best of our knowledge. The novelty of our work lies in providing solutions for these unsolved problems.

We present solution for identifying palmprints having 3D perspective projections and having a variety of illumination and contrast variations. We now present the existing literature for solving pose variance and handling illumination variations. We then analyze the applicability of those solutions to our problem domain. A detailed study and discussion has been provided below.

There are various techniques used for the matching of objects in different poses. The objects themselves, fall in two categories: a) non-deformable and b) dynamic. *Non-Deformable objects* have a fixed shape and form, which gives them a fixed structure. For eg. a ball, table, chair etc. are non-deformable objects. While modeling these objects from different views, a knowledge about their fixed structure helps in extrapolating points in a different pose. Dynamic objects, on the other hand, are the ones having no fixed shape and form. It can be characterized in two categories: a)Articulated objects and b)Non-Rigid(deformable) objects. *Articulated objects* are those that can be divided into rigid objects connected by joints; for eg. biological forms having a bone structure viz human body parts like limbs, hands etc. *Deformable/Non Rigid objects* comprise of the objects whose structure changes with time. Examples include a swimming jelly fish, wheat stalks blowing in the wind, and a beating heart. Deformable objects in different poses need more intricate solutions as the structure of the object is not completely predictable. A brief overview of the state-of-the art techniques being used is given below.

Pose Variation in Non-Deformable Objects

The solutions proposed in literature for solving pose variations of non-deformable objects can be classified into the following main categories [76]:

Fiducial Based Matching: In this type of scene matching, fiducials, also called landmarks or markers, are bright external objects inserted into the scene or are placed around the object. For eg, easily recognizable bright red props are strewn in the scene. These markers are easily recognizable in an image and form correspondence points from various different views. The motion parameters can then be computed based on the shift in the position of these fiducials and then applies to the entire scene to get the correct position of the rest of the objects. Fiducials can be both circular ([77]) and planar([78]). The circular fiducials give a point match whereas the corners of a planar fiducial provide correspondences between an entire plane. These fiducials may vary in sizes ([79]) and colours ([80]).

Model Based Object Matching: Although reliable, the insertion of physical markers in the scene is not a very good solution to the problem of finding point correspondences in objects captured in different views. It is hence preferred to rely on the features naturally present in the image. For eg. line features, planar parts or a knowledge of shapes can be used to detect those in the image. This makes the point tracking problem much more challenging and more practically apt. Model based object matching is an iterative procedure which needs initialization. The point correspondences initialized once (manually or in an *ad hoc* manner), are used to find out the motion parameters initially and then to model the motion by predicting the point locations in the coming frames to help look for image features. However, minor errors in pixel position cannot be avoided in this kind of an approach. Also, choosing the approach to use depends on the application at hand. It is required to have some a priori knowledge about the object/scene for using model based tracking. The information used in the images can broadly be classified as edge based and pixel

based. A brief overview of these has been provided below.

Edge Based Modeling: The edge information present in the image can be reliably detected in different scenes. Interest points can be found by sampling these images at along the edges at certain frequencies. However, finding correspondence between these points can be tricky. This is because of the errors in pixel position that are introduced due to the sampling of the edge. The advantages of using this approach are computational efficiency (because of considering only edge pixels rather than the entire image) and partial invariance to the illumination variation (the robustness of the invariance depends on the edge extraction algorithm used). Edge based methods can be further grouped into the following two categories:

- One can look for strong edge responses present near to the location predicted by the pose model computed in the previous iteration. This does not need an explicit extraction of the contours present in the image. ([81, 82, 83, 84, 85]). This is fast and a generic approach which does not exploit the advantage of the information present in the image of the new pose.
- Another approach is to explicitly extract the image contours, such as straight lines and boundaries and then recompute the model to fit these contours ([86, 87, 88, 89, 90]). The loss in generality can be compensated with a gain in robustness.

Optical Flow-Based Matching: Optical flow is the apparent motion of the projection of a physical point in an image sequence, where the velocity at each pixel location is computed under the assumption that projections intensity remains constant. While solving for pose variation, flow can be computed using two images only. It can be expressed as :

$$m' = m + \begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} dt \quad (2.3)$$

It can be computed using the Lucas-Kanade method [91] for example, which adopts a multi-scale approach and assumes that the optical flow varies smoothly. Slight Variations of the optical flow technique have been proposed in [92, 93]. Optical flow and edge information can also be combined to find out the motion of the interest points on the edges to avoid error accumulation [94, 95].

Template Matching: The Lucas-Kanade algorithm [91] was originally designed to compute the optical flow at certain pixels of interest in an image, but it can be used in a more general sense to register 2D templates. Registration is the process used to align two images, which also needs to compute the transformation parameters in order to transform the image[91]. The Lucas-Kanade algorithm when used for image registration does not rely only on local features such as edges and interest points. It uses the entire pattern of the image in a global sense to model complex objects which cannot be defined using local features only. Since it uses the complete image pattern, it is computationally expensive, but [96] showed that using certain constraints and special conditions, it can be effectively formulated. After that, it has been tried for 3D tracking as well by several authors [97, 98].

Interest Point Based Methods: After looking at algorithms using global features, we now consider algorithms using local features, just like the edge based modeling. Extracting localized features has several advantages over global features. Since, we look for individual features, occlusions(when the object is only partially visible in the image), and illumination invariance can

be achieved to some extent. The advantage over edge based modeling is that it does not mix up background boundaries with features of interest and does not have to rely on the accuracy of the edge detection methods. Their advantage over optical flow methods is that they can allow for faster and non-uniform motion in the image pixels.

These localized patches in the image can either be initially specified by the user, or be automatically computed by the algorithm. In early works, pixels with minimum variance in intensity in the four directions were retained as useful patches [99]. The currently popular detectors, i.e. the Forstner operator [100], the Plessey operator, the Harris-Stephen detector [101], or the Shi-Tomasi [77] detector [102], all rely on correlation matrices computed at each pixel location. Eigen vectors can also be used to classify image points into useful features [100]. Pixels with two large, approximately equal eigen values are good candidates for selection.

Finding Point Correspondences by Detection: The iterative nature of all the pose detection methods mentioned above estimate the pose parameters of the next iteration beforehand and then optimizes the error in computation. However, this approach always needs to be initialized and the results are affected by the accuracy of the initialization. Pose estimation from naturally occurring image features without prior on the actual position is closely related to object detection and recognition. This problem has been studied in computer vision domain for long for both 2D and 3D objects [103, 104].

Early approaches were edge-based [105, 106], but methods based on feature points matching have become popular since [107] which shows that local invariants work better than raw patches for such purpose. [107] uses invariants based on rotation invariant combination of image derivatives but other local invariants have been proposed later. It was realized that feature point based methods appear to be a better approach to achieve robustness to scale, viewpoint, illumination changes and partial occlusions than edge- or eigen-image- based techniques [55].

The interest points once detected in the image are described using *local descriptors*. Local descriptors basically imbibe the context of the neighbourhood resulting in a feature vector at each interest point. Many such descriptors have been proposed over the years. For example, [107] computes rotation invariant descriptors as functions of relatively high order image derivatives to achieve orientation invariance; [108] fits an ellipse to the texture around local intensity extrema and uses the Generalized Color Moments [109] as a descriptor. [110] introduces a descriptor called SIFT based on multiple orientation histograms, which tolerates significant local deformations. This last descriptor has been shown in [111] to be one of the most efficient. SIFT is one of the most popular descriptors used for object detection and matching. It is invariant to scale, rotation and occlusions. Every interest point is assigned a feature vector of 128 dimensions having information about the point neighbourhood. It is computationally efficient as well.

Pose Variation in Deformable Objects

There has not been much literature providing an overview of the existing work on solving pose variations of the deformable objects. However, many researchers categorize it in the following two broad classes. 1. Active Contour based methods, and 2. Model based methods and. These have been described in more detail here.

Active Contour Based Methods: Active Contour model or snakes is a method to locate contours or boundaries in an image [112]. This is done by fitting a 2D elastic model consisting of splines (piecewise polynomial) to the roughly estimated initial boundary points. The spline then

grows to minimize an energy function and fits on the boundary after a few iterations. These models can be used to find out the change in the image pose by evaluating the change from one contour to another. In [113], et al used active contours to locate facial features in different images. The splines were designed to fit the facial features. This could be used for modeling the palm lines. However, the palm lines are generalized better as a combination of straight lines, and designing a polynomial to fit them would probably only increase computational complexity. Additionally, the problem of finding the initial contour or palm line accurately enough for spline generations remains challenging as even small breaks can disrupt the process of contour initialization.

Model Based Methods: Metaxas ([114]) proposed a 3D mesh governed by the laws of Physics for elastic bodies to identify motion in both rigid and non rigid objects. Following a set of rules which also govern the actual motion of the object results in better accuracies than obtained with active contours using energy functions. Two types of generic models can be used for this kind of modeling, namely, Finite Element Modeling(FEM) and Boundary Element Modeling (BEM). As the name suggests, Boundary Element Modeling fits the mesh on the boundary of the elastic object as opposed to the FEM approach which requires the object body as well as the boundary to be meshed.

2.4.1 Solving Pose Variation for the Palmprint

In our work, we consider images taken with unrestricted and mostly unsupervised hand motion. This may lead to a change in hand pose. Typical set of images to be matched from our dataset are shown in Figure 3.1. The free mobility of the hand and the fingers result in many creases on the surface of the palm which cause a change in appearance of the texture. Even when holding one's palm in a stretched out manner, one notices different creases as compared with a relaxed palm. Due to this complicated inter-related motion of fingers and the palm skin, false line-like impressions are created on the palm. Hence, the palm can be characterized as a combination of articulated and non-rigid object.

When the users are asked to pose for the camera with a stretched out palm, some of the intricacies in palm shape mentioned above can be done away with. The non-rigidity induced by skin creases can be ignored. The palm can then be studied as an articulated object with different rigid parts formed by the finger to palm joints and a thumb to palm joint. Since we do not capture the information present in fingers here, we focus only on the central region.

We can approximate the palm to a rigid object when it is posed in a stretched out fashion. It is actually made up of two rigid parts, first a small elevated ridge-like region connected to the thumb and another plateau-like region which comprises of the major portion of the palm. The ridge-like region can be modeled using a complex polynomial whereas the relatively flat plateau is appropriately represented by a plane. In most cases, changes in one do not affect the appearance of the other part unless intentionally done so by a non-cooperative user.

Hence, for the sake of simplicity, we approximate the entire palm as a plane for the purpose of solving for pose variation. This is acceptable since the plateau comprises the major part of the palm and the height variation of the ridge is not considerable. We do note that the recognition accuracy can be improved by modeling both these regions separately and then solving for pose parameters. However, whether the associated increase in complexity is justified by the increase in accuracy remains to be verified.

To sum up, we capture palm images using a generic web camera in a completely unconstrained

and unsupervised environment. The only restriction that we impose is that the user's hand be stretched out while posing for the camera. Under this condition, we can approximate the palm region with a rigid plane. We implicitly choose algorithms to imbibe the following peculiar characteristics of this rigid planar surface which are challenging to solve for:

- **Low-textured image.** Palmprint is a lowly textured image. The only texture present on the palm image is in the form of palm lines. Please note that although this texture suffices for recognition purpose, it is certainly not rich enough to extract feature points to find point correspondences to solve for pose variation.
- **Illumination Variations.** The images we consider here suffer from a variation in illumination variation and noise features and hence, making the extraction of reliable points extremely difficult.
- **Wrinkles** present on the palm due to the looseness of the palm skin need to be acknowledged when using an edge detector. We should be able to robustly differentiate major features from noisy line-like features.

Hence, we design a technique for pose invariance keeping the above challenges in mind. Edge based tracking suffers in the presence of illumination variation and background variation. Since we capture generic images with no constraints on lighting or background, it cannot be used[55]. Optical flow based techniques cannot be used because the motion of the hand could be swift(resulting in two very different poses of the hand). Optical flow cannot handle a large variation in pose transformation parameters. Hence, when the hand is shown in two very different poses, it fails to find suitable correspondences.

When considering planar textured objects in different poses, without occlusions, template based matching is reported to perform accurately [97]. It is unfortunately very sensitive to occlusions, illumination and scale variance. Direct Template Matching is more appropriate for small planar objects having a defined geometry like books, tables etc. In contrast, interest point matching does not suffer from occlusions. It can be made invariant to illumination and scale depending on the robustness of point detection and selection process (for eg. SIFT). However, it needs a highly textured object for reliable point detection. This is not possible for the lowly textured palm image. To give an idea of the difference in texture of palmprint compared to other objects, one can consider the SIFT features found by the algorithm on the palm compared to any other object. As compared to thousands of features matches normally found, SIFT found to detect only a handful(around 6 to 7) corresponding features on the palm.

Hence, we must design innovative ways to find reliable points on the palmprint to solve for the pose variation problem. A robust solution to this has been provided in the next chapter.

2.4.2 Illumination variations

Much of the work focused on illumination invariant biometric recognition has been focussed on face recognition. This is the biometric modality which is easy to capture using a generic camera and is highly user friendly. A related survey of popular methods can be found in [115]. Much effort has been made to model the illumination on different parts of skin and the various changes due to face pose. A Lambertian model without shadow was proposed in [116]. Georghiades *et al* [117]

proved that face images having the same pose and different illumination settings fall under a convex cone. This cone of variations has been termed the illumination cone. Different representation using the quotient images [116], and spherical harmonic representation [118, 119], were proposed which were invariant to illumination. The use of infra red light in place of the visible light was also proposed[115].

It is clear that one of the approaches used for achieving illumination invariance is to model the illumination variation or to estimate the reflectance model of the light from the face [120]. Both these approaches are mathematically intensive and it is not always possible to model the surrounding light. Some other direct approaches to achieve partial illumination invariance for palmprint recognition was used in [8]. The mean and variance of the image is set to a prefixed value and hence, restricting the impact of uncontrolled lighting to some extent. However, this technique was found to blur out the line information present in the image which was especially hazardous in the case of varying palm poses. This was due to the fact that the lines in different poses have very less overlapping information to begin with.

Another method to avoid illumination variations for palmprint recognition suggests the use of imaging using infra red light based cameras, which is able to capture a relatively more stable palm image. But this would require us to add additional hardware and is not possible using generic cameras. We need to solve the problem of illumination variation in images captured using a commodity camera. Hence, we look for new methods to solve the problem of illumination variation for palmprint recognition. A solution which does not need explicit modeling of the illumination and hence can be taken care of dynamically on the run without any prior training of the illumination model.

2.5 Summary

In this chapter, we discussed about the state of the art palmprint recognition process. The importance of pre-processing of the image to extract the palm image in the recognition process is apparent. With a change in the image capture process, many of the current approaches are no longer valid. Image matching stage also needs to be modified to accommodate the various artifacts present in the image. Before matching, the palm must be corrected for pose transformations. The feature matching stage should be robust to illumination variations. The existing approaches for pose correction are developed mainly for object recognition. As compared to palmprint, most objects have a richer texture quality, which is an asset for finding point correspondence. Hence, we need to look for new approaches to find point correspondences in lowly textured images. Illumination variations is also an unsolved problem. The only existing solution to this problem, for the case of palmprint, is to fix a mean and a variance of the image intensity values to a pre-fixed value. This is not enough for extreme lighting variations. Other solutions are complicated for use in a real-time environment. Thus, we need to find solutions that work well for unconstrained palmprint imaging accounting for the low texture of the palm and the variable nature of skin. We describe the solutions to these problems in the next two chapters. Chapter 3 proposes a novel method to solve the pose variation problem and Chapter 4 describes an approach to address illumination invariance.

Chapter 3

Pose Invariant Palmprint Recognition

A palmprint based authentication system that can work with a multi-purpose camera in uncontrolled circumstances, such as those mounted on a laptop, mobile device or those for surveillance, can dramatically increase the applicability of such a system. However, the performance of existing techniques for palmprint authentication fall considerably, when the camera is not aligned with the surface of the palm. The problems arise primarily due to variations in appearance introduced due to varying pose, but is compounded by specularities of the skin and blur due to motion and focus. In this chapter, we propose a method to deal with variations in pose in unconstrained palmprint imaging. The method can robustly estimate and correct variations in pose, and compute a similarity measure between the corrected test image and a reference image. Experimental results on a set of 100 user's palms captured at varying poses show a reduction in Equal Error Rate from 22.4% to 8.7% (Section 3.6).

3.1 Introduction

The use of digital imaging, specifically camera based imaging, for biometric authentication has changed the applicability of biometric authentication for personal security. Camera based imaging is fast, inexpensive, convenient to use, and is easily available to the common man. Possibly one of the biggest advantages of camera based imaging is that it allows easy image capture in uncontrolled and natural settings, a feature prominently missing in the traditional ink based prints or scanners. If used properly, this makes any authentication system very convenient and intuitive to use. In addition to biometric modalities such as face, gait, etc., one can employ camera based imaging for fingerprints, palmprints, or even handwriting. This facilitates the use of multi-modal biometrics, which alleviates many concerns on accuracy and applicability, as it is possible to capture different biometric features using the same imaging setup and at the required resolution.

But despite all the advantages, there are many challenges to overcome to realize biometric authentication systems that use unconstrained camera-based imaging. The major problems include background clutter or noise that make segmentation difficult, illumination changes that can affect the feature extraction process, scale and resolution changes, and the increased flexibility offered to the subject in terms of pose changes (Refer Chapter 2). The onus then lies on the algorithm to accommodate all these variations in the best possible way. This adversely affects the recognition rate which is much higher when limited degree of freedom is allowed.

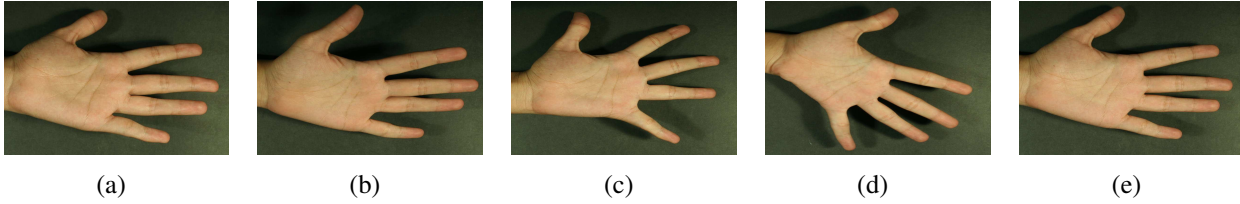


Figure 3.1: Variations in hand pose due to unconstrained imaging.

As a result, many camera based recognition systems employ a closed box-like capturing setup with pegs on a supporting board that constrains the pose and scale, and provides uniform illumination and constant background ([8],[121])(See Figure 2.5(a)). Such a setup may not add much in terms of cost, but makes the apparatus cumbersome to use, liable to frequent damage, less intuitive and user friendly, and very specific to a particular biometric. Thus, the major advantages of using a generic camera remain under utilized. Hence, there is a potential to develop new algorithms that are more robust to environmental variations and also allow more flexibility to the subject. Efforts in this direction include achieving pose and illumination invariance for face recognition [18], view invariance for gait recognition [28] and handling different poses for hand geometry based authentication [122].

In this chapter, we focus on the problem of achieving pose invariance in palmprint based authentication. The hand is the most flexible organ in human body with high degree of movability. It makes it highly convenient to use hand based authentication both as a single modality and also as a supplement to another biometric modality in a multi biometric system. The high degree of freedom of hand also results in a variety of poses which makes it necessary to deal with pose variation to develop a robust palmprint recognition system. However, the performance of existing techniques for palmprint authentication fall considerably, when the camera is not aligned with the surface of the palm. The problems arise primarily due to variations in appearance introduced due to varying pose, but is compounded by specularities of the skin and blur due to motion and focus. To solve the problem of pose invariance in palmprint recognition, one must also partially address all these problems, since finding palm lines accurately is necessary for any palm based recognition algorithm. We present a novel solution for this problem in this chapter.

The rest of the chapter is organized as follows. Section 3.2 discusses the related work in this area. We then describe the novelty and contribution of the work in Section 3.2. Section 3.3 explains the approach chosen to attain pose invariance and reasons for it. Section 3.4 describes the practical challenges in image alignment and the solution. Section 3.5 gives a step wise formulation of the solution proposed by us, followed by experimental results in Section 3.6.

3.2 Previous Work

There are many unique features in a palmprint image that can be used for personal identification. Principal lines, wrinkles, ridges, minutiae points, singular points, and texture are regarded as useful features for palmprint representation([123]). Different features can be extracted at different image resolutions([8]). For finer features such as minutiae points, ridges, and singular points, a high-resolution image, with at least 500 dpi(dots per inch) ([41]), is required for feature

extraction([124]). On the other hand, the coarse features such as the principal lines and wrinkles, as shown in the figure, can be obtained using a low resolution image,(less than 100 dpi ([125]),[123]). Since, a high resolution image gives the coarse features as well as the fine features, the recognition accuracies are generally better for high resolution images. So, the choice of the fineness of the feature depends mainly on the application for which palmprint recognition is being used. For applications such as law enforcement, the confidence in outcome is more important than computation time and hence finer features such as ridges, singular points, and minutiae points are used for palm matching. For civilian and commercial applications, low-resolution images are more suitable. Since these low resolution images have smaller sizes, the preprocessing and feature computation processes take lesser time and hence making real-time computation possible. We shall now take a look at the methods used for extracting both the finer and the coarse features for a given palmprint.

Initial palmprint matching systems used images generated by digitally scanning inked palms. A detailed image having the major lines, wrinkles and creases of the palm was obtained. Approaches developed for these images involved matching datum points, and the line orientation at them ([39],[125]). The detail and clarity in images obtained by inked palms is such that some researchers even used the palm creases to carry out the identification process [126]. This constituted the offline palmprint matching. An online system captures palmprint images using a palmprint capture sensor that is directly connected to a computer for real-time processing([8]). Real-Time Online palmprint matching was made possible with the use of electronic imaging by capturing low resolution images using devices such as web cameras. The setup is usually designed so as to restrict the hand motion and the surrounding effects on the image. The subjects are required to place their hand on a planar base, which helps in fixing the orientation of the hand. Further restriction is provided by fixing pegs on the planar base, such that all the fingers are placed at a pre fixed position ([47],[127], [128],[39],[125]).

The assumptions of controlled environment and restricted palm pose were convenient, and only prominent palm lines and their structure were used for matching ([8],[129]) resulting in high accuracies. As both hand-geometry and palmprint use the same imaging setup, one could also combine them both in a multi-biometric setup [121]. The common preprocessing step computes the two valley points p_1 and p_2 as shown in Figure 3.2. As the palm is cut out with respect to the coordinate axes formed by joining these two stable points on the palm, the extraction process becomes invariant to in-plane rotations and translation. Hence, in recent past, ongoing research on palmprint recognition has mainly been focused on finding better and more effective ways of feature representation for improved and robust matching. As reported in [130], the reported efforts to find suitable features for palmprint recognition can be classified in the following three categories: a) texture-based approaches, ([8], [131], [132], [133]) b) line-based approaches ([121], [134], [123], [129]) and c) appearance based approaches ([128], [127]).

All the approaches mentioned above have reported considerably high recognition accuracy. This is due to the partial rotation and translation invariance to in-plane rotation and translation, achieved by constrained pose hand imaging. But methods used to control illumination generally increase the volume of the system and having to touch the fixed plate while posing for camera is considered unhygienic by some. It also suffers from problems relating to user inconvenience, being less intuitive, unsatisfactory user cooperation, a cumbersome setup and limiting the possibilities of extending the system. These issues must be addressed to make the system more robust and better adept to practical scenarios.

Research is underway to develop touchless or contactless palmprint recognition system ([36],

[37], [38], [74]). (Figure 2.5(b))The system deployed is reported to use a fixed camera to capture hand image. They got rid of the fixed board, that was previously used to support the subject’s hand, which was a hygienic concern for many. With the removal of the board, the pegs used to fix the hand in position, which were looked upon as a source of discomfort, were also discarded. The use of such a system is the first attempt in the direction of making the palmprint image capture a more user friendly experience. The subjects are however, required to place their hand parallel to the imaging plane of the camera, and the lighting conditions are either artificial or controlled. Hence, these systems make assumptions pertaining to user co-operation and the availability of controlled lighting conditions which plays a very important role in the line extraction process, which shall be shown later.

But the subjects may not place their hands parallel to the camera’s imaging plane. These pose variations will highly affect system’s recognition accuracy. Thus, the problem of achieving a truly unconstrained palmprint recognition process which can handle pose variations and variable illumination settings is still unsolved even though the issue of achieving pose invariance for hand geometry based recognition has been attempted. A system for identifying individuals based on hand geometry signatures, using an unconstrained imaging device, was developed by G. Zheng *et al*[122]. The subjects are not restricted to pose their hand parallel to the plane of the camera. To handle these cases, Zheng *et al.* created projective invariant features from the hand. This approach, however, can not be used to solve pose invariance of palmprint recognition. The projective invariant features used in [122], are a group of cross ratio values, which are defined over a set of fixed points found on the hand. To be more specific, these points are taken from the creases found on fingers. Since, it is very difficult to isolate credible landmark points on the palmprint, it is difficult to find projective invariants when comparing two different palm poses. Solving this will, no doubt, enhance the recognition accuracy achieved in [122], and make an effective and robust pose invariant multi-biometric hand biometric recognizer.

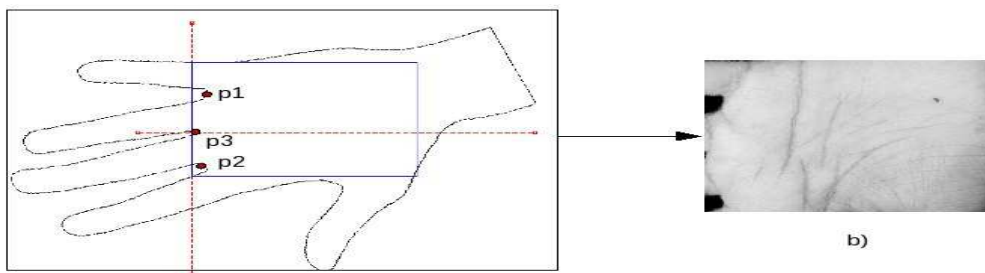


Figure 3.2: Extraction of palm region based on finger valleys.

In this work, we extend contactless palmprint recognition to include matching images with hand variations in 3D space(see Figure 3.1) and non-restricted illumination, making palmprint recognition possible using a truly unconstrained imaging setup. Pose variations distort the absolute position and relative distance between the palm lines, and change their characteristic properties like length, width, curvature, aspect ratio etc. Hence, traditional matching can not be applied directly to palm images having different poses. The variable lighting makes the process of extracting palm lines extremely difficult. Hence, we need to look at novel methods of palm line extraction to

solve for this.

Novelty And Contribution

In this work, we approach the problem of unconstrained palmprint imaging and the related pose variations for the first time. We present the entire framework that needs to be used to enable unrestricted palmprint imaging and discuss novel solutions for the yet unsolved bottlenecks. We propose a novel algorithm for pose invariant palmprint matching. As a part of the algorithm, we propose a robust image alignment technique and a method to detect the orientation of the hand. For demonstrating the results of the pose invariance algorithm, we created two datasets: a synthetic and a real dataset. The synthetic dataset gives us an opportunity to isolate, quantify and incrementally study the effects of pose changes, while the real dataset accurately models the practical problems such as skin deformations, specular reflections, etc. Some images from our dataset are shown in Figure 3.1.

The rest of the chapter is organized as follows. Section 3.3 describes the challenges faced in making an unconstrained imaging based palmprint recognition system. Section 3.4 proposes our solution for achieving pose invariance. Section 3.6 gives a description of the datasets used, followed by the experimental results and discussions.

3.3 Unconstrained Palmprint Imaging

In our work, we consider images taken with unrestricted and mostly unsupervised hand motion. This may lead to a change in hand pose. Typical set of images to be matched is shown in Figure 3.1 (a) and 3.1(b). Matching can be done in two ways.

3.3.1 Finding projective invariant features in palm

As pointed out in [122], when subjected to geometrical projective transformations, cross ratios created by 5 coplanar points with a specific configuration are invariant to pose change. Each five point set results in 30 feature values depending on the possible permutations of the arrangement. The feature vector is created by using 12 landmark points to generate $C_5^{12} = 792$ groups of points, each resulting in 30 features. The final feature vector has a length of $792 \times 30 = 23,760$. The approach assumes the planarity of hand, and the landmark points are chosen from the the finger creases.

The bottleneck of this is that it requires us to find 12 stable landmarks with specific geometric relationship on the palm images. However, considering the structure of palm, which consists of three major lines: the heart, head and life lines, the number of distinctive points that can be reliably identified are very limited. Points formed by intersection of two lines are not reliably detected due to blur, illumination and resolution changes, and the problem increases further due to introduction of wrinkles during pose changes. In short, the palm images are weakly textured and do not have enough acceptable landmark points to consider the computation of projective invariant features.

3.3.2 Pose correction

A second approach is that of explicitly determining the pose and then correcting for it. A fundamental matrix generally models the transformation between objects in 3D world

$$[x', y', c] \mathcal{F}[x, y, 1]^T = 0, \quad (3.1)$$

However, we can restrict the set of possible transformations, by assuming the thickness of the palm to be negligible, approximating it to a planar surface. Considering the palm to be planar is reasonable and holds true in most cases[122]. It should be noted here that the palm, when in different poses, becomes prone to deformations of the skin because of the movement of the fingers and the thumb. Hence, additional line like impressions are introduced. These unwanted lines can be controlled to an extent by using certain line-extraction techniques, which will be described later in this section. However, we do not focus on explicitly discarding wrinkles and other noise in this section. It has been separately addressed in the next section. Hence, for all practical purposes, we assume the palm to be a rigid plane, thus causing correction of pose variation to become well posed. From multi-view geometry of imaging, two images of a planar object are always related by a linear transform, characterized by 8 parameters [135]. Given any point in one image, in the homogeneous coordinate system, $[x, y, 1]^T$, the corresponding point in the second image is given by:

$$[x', y', c] = \mathcal{H}[x, y, 1]^T, \quad (3.2)$$

where the actual image points are x'/c and y'/c . The homography matrix, \mathcal{H} , can be estimated from correspondences between two views, up to scale. Hence \mathcal{H} has 8 unknowns and we need 4 point correspondences to solve it.

But, as mentioned earlier, it is difficult to find robust correspondence points in the weakly textured palm image. So, its natural to find a solution that needs as little point correspondences as possible. On comparing the two approaches, we can see that correcting for pose requires 4 reliable point correspondences, whereas for creating projective invariant features we need 5 or more reliable correspondences. Here, we would like to stress on the need to find reliable point correspondences, and reiterate the fact that it is not possible to find reliable correspondences from palm lines. At best, we can find probable matches, since it is highly difficult to tell which point corresponds to which when comparing two lines with no distinctive curvature properties. Hence, the solution must be such that it takes a set of these probable matches and refines them to result in a set with reasonably high confidence. RANSAC based homography estimation is one such procedure [136]. The flexibility offered by RANSAC in choosing the initial point set, makes it possible to proceed with a set probable correspondences. The problem is now reduced to finding this set of probable matches before proceeding for the refinement and the pose correction step. We call this correction the Image Alignment stage, which is described below. (For details on RANSAC Homography, please refer to Appendix)

3.4 Image Alignment

Our goal is to align two palm images with each other so as to match them properly. For this, we need to find 4 point correspondences for applying Homography for pose correction. In the absence

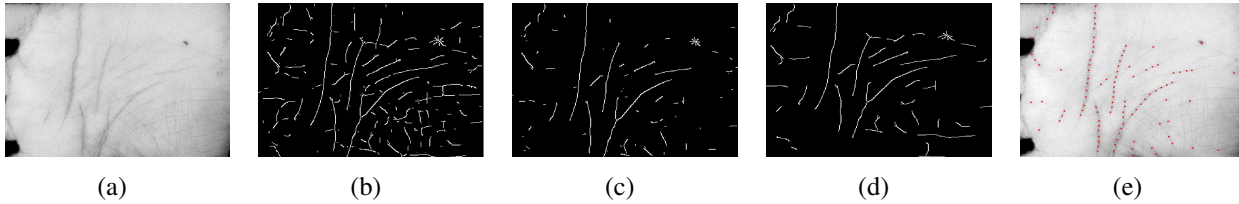


Figure 3.3: The process being followed to get interest points from the image. a) Palm Image, b) Edge map with a lower threshold, c) Edge map with a higher threshold, d) Edge map after enhancing image in c, e) Sampled interest points marked on the palm image.

of robust correspondences, we use a set of interest points as the initial correspondence and try to improve it using RANSAC. The most stable points on the palm are the valley points, p_1 , p_2 and p_3 shown in the Figure 3.2, and we directly know the correspondence between them. It is not possible to use all three points for homography, since these points are almost collinear. Collinear points do not form good correspondences for Homography[135]. Hence, we can only use two of these points, say p_1 and p_2 . For the other two correspondences required, we search the palm lines for any points with a strong surrounding context in both images. The reason for this being that although the global context can help us find the first set of interest points, it does not offer a lot of distinguishing information to differentiate points located close by on the same line. The local context around the points must be used to differentiate between the points on the same line. The points with high confidence will be the ones that are formed by line-line intersection, line-wrinkle intersection or maybe wrinkle-wrinkle intersection. Points where branching of lines occurs, resulting in Y-shaped structures have distinctive context around them, making it possible to find their correspondences. Sometimes, there may be no such unique point available to us depending on the line pattern, in which case, we need to have a backup solution.

The process of n set of probable matches, and call them interest points. From this set, we shortlist the required corresponding points. The process of interest point detection proceeds as follows: We find an edge map of the lines using a curvature based method proposed in [137]. The approach is more robust to noise, and extracts prominent lines in the image. However, in many cases the edge map is still noisy. So, we adopt a two-step line detection algorithm. First, two edge images are computed with a high and a low threshold(Figure 3.3(b) and (c)). The image with higher threshold has parts of all the major lines and prominent wrinkles. We use these points as seeds to search in the low-threshold image using morphological operations to get the final edge map. The final image obtained is shown in Figure 3.3(d). In the second step, we richly sample points from edge map as indicated by the cross marks in Figure 3.3(e). These randomly chosen points form the initial candidates that have the highest probability of being the interest points. The initial point set is culled by removing those without significant correspondence in the second image. The correspondence level is estimated by using correlation(See Algorithm 1). The list of corresponding points selected by this process is referred to as putative matches.

There are possible scenarios where the correspondences found are incorrect. Firstly, there will be cases where there are no intersection points or points with distinctive local context, or when those points are not selected in the random sampling, then we choose the most similar looking points. Secondly, there are possible scenarios where the correspondences found are incorrect, since we have used the local context only. Another possibility is the introduction of additional

interest points due to skin deformation or illumination difference. To compute a reliable estimate of homography with these putative matches, we use a RANSAC based method [136]. The input to RANSAC is n pairs of putative matches, including the 2 stable points computed from the finger valleys. We modify the traditional RANSAC algorithm slightly to always include the two valley points in the set of 4 points, as we have high confidence in them.

Finally, we transform the test image with the computed Homography matrix. This results in the sample image being modified as per the dimensions of the template. Homography computation is sensitive to even small errors in point correspondences. The efficacy of the entire algorithm finally depends on how good correspondences turn out to be. The overall process of image alignment is presented in Algorithm 1. A typical set of input images and the aligned output image are shown in Figure 3.4.

3.5 Proposed solution

The flowchart of the state-of-the-art palmprint recognition process has been given in Figure 3.5(a). To make it pose invariant, we need to introduce the image alignment step(Figure 3.5(b)). This can not be achieved without slightly altering all the remaining steps of the algorithm. Since the images are captured in an unconstrained setup, the preprocessing and the feature extraction steps need to be modified according to the image alignment step. The overall matching algorithm includes the following steps: preprocessing, extraction of the region of interest (ROI), image alignment, feature extraction, and matching.

3.5.1 Preprocessing

The preprocessing step needs to be altered the most in this case. It consists of four steps, a) Extracting hand from the background b) Binarization, c) Determining the orientation of the hand, and d) Locating the two finger valleys.

Extracting Hand from Background: To segment out the hand from a cluttered background, an ada boosting based approach was proposed by Doublet *et al.* [74]. Stenger *et al.* used skin colour and hand shape to extract out the human hand from a variety of different backgrounds. However, this was demonstrated on videos of hand ([71]) and exploited the continuity of the hand motion of the video sequence to extract the hand image. Hence, this may not be directly applicable to images. Hence, extracting hand shape from a cluttered background is a non-trivial problem domain in itself. In this work, we use a uniform background in our experiments as our focus is to

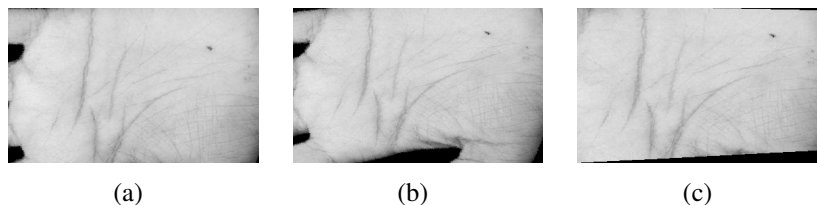


Figure 3.4: Result of aligning Image a to Image b, using Homography, has been shown in Figure c.

Algorithm 1 Algorithm for Image Alignment

Input: $Image_1, Image_2$ with different poses.

Output: *pose corrected Image₂*.

1. Compute interest points in each image by richly sampling points on the edge map.
 2. Use correlation, with a window size of 11, to compute the putative correspondences. If the number of correspondences obtained is less than some threshold, label it as imposter.
 3. *RANSAC based robust Homography estimation:*
 - (a) **for** $k = 1$ to N **do**
 - Selection of Point Set:** Select the two point pairs randomly from the putative matches and compute the Homography H_k by including the two valley points.
 - Distance Computation:** Calculate the distance d_i for each pair of putative correspondences after Homography transformation.
 - Compute Inliers:** Given a threshold t , compute the total number of inliers, n_k among putative matches, such that $d_i < t$.
 - (b) **end for**
 - (c) Choose the final Homography, H as $\arg \max_{H_k}(n_k)$.
 4. *Output the image obtained by transforming Image₂ with H.*
-

solve the problem of unknown hand pose for matching. The actual segmentation process is then carried out along with the binarization step explained below. We redirect the reader to refer to the works mentioned above to get a detailed solution of the problem of extracting hand from a cluttered background.

Binarization: In the approaches mentioned for the state-of-the-art Palmprint Recognition in Chapter 2, we observed that thresholding is a pretty straight forward procedure. Some methods used a fixed threshold[8], while others used Otsu’s method to generate a dynamic threshold[121]. But since the background and the illumination settings in the image capture step are fixed and uniform through images, clean thresholds are obtained and the hand gets clearly segmented out. We capture images in an uncontrolled and mostly unsupervised manner. Since the lighting conditions in our setup are relatively uncontrolled, binarization becomes difficult. Straightforward thresholding in intensity domain does not work. It results in broken patches. So, we perform binarization by thresholding using the hue of the image instead of intensity. Intensity captures the ‘brightness’ of the image whereas hue measures the ‘pure colour’ content of the image[138]. Illumination variation in the scene mostly affects the brightness of the scene. The changes in hue due to illumination variation are relatively smaller. Hence, the binarization process is carried out by first converting the image from RGB (Red, Blue, Green) space to the HSI (Hue, Saturation and Intensity) space, and then thresholding the image based on hue rather than the intensity. This works fine in our case since we have only two colour classes namely the constant background and the hand. Therefore,

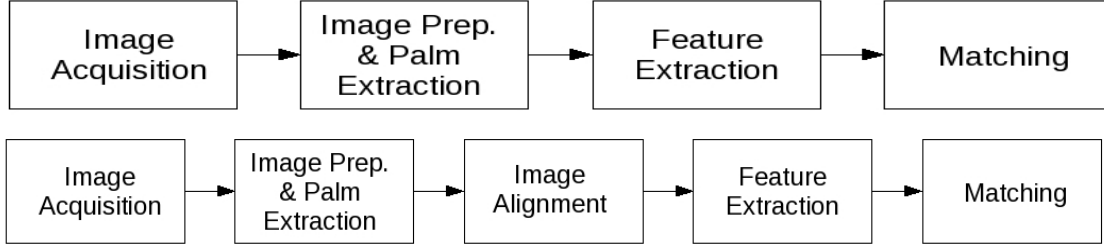


Figure 3.5: a) Flowchart of state of the art palmprint recognition algorithm. b) Proposed algorithm.

thresholding results in a clean segmentation of the hand from the background.

Orientation of the hand: In most of the processes seen in the state-of-the-art Palmprint Recognition process seen in Chapter 2, hand orientation is either fixed apriori or determined with the help of pegs. Some methods find the orientation by model fitting (ellipse) on the hand and taking the major axis to be the dominant orientation. In our approach, we do not fit any model on the hand. We find the direction of the maximum 'spatial variance' in the foreground (which is the hand shape) and fix that as the hand orientation. This can be done by creating a scatter matrix for the binarized hand image and then finding out the largest Eigen Vector of this scatter matrix. If I is the binarized palm image with dimensions $m \times n$, then we define a matrix $M = [X Y]$, where

$$X = x_i, Y = y_i \mid 1 \leq x_i \leq m, 1 \leq y_i \leq n \ \& \ I[x_i, y_i] \neq 0 \quad (3.3)$$

The scatter matrix for the above matrix M can be defined as:

$$S_x = x_i - m_x, S_y = y_i - m_y \text{ s.t. } m_x = 1 \div \text{sizeof}(X) \times \sum_{i=1}^m x_i, m_y = 1 \div \text{sizeof}(Y) \sum_{i=1}^n y_i, x_i \in X, y_i \in Y \quad (3.4)$$

The largest eigen value of this scatter matrix, S , of M above gives us the orientation of the palm. This method corrects large degrees of in-plane rotations as opposed to previous approaches.

Finally, we extract the valley points as in [8]. The result has been shown in Figure 3.2.

3.5.2 Extraction of ROI

The region of interest in the hand image is defined with respect to the co-ordinate axis formed parallel to the line joining the two valley points, p_1 and p_2 detected above. We assume the middle valley point, p_3 to be the origin (Figure 3.2). We then extract the maximal rectangular region aligned to the above line, containing foreground pixels. This allows us to consistently extract the entire visible area of the palm, irrespective of the change in view. Note that this is an approximate region of interest, and the matching is done after image alignment.

The palm extraction process indirectly solves the problem of variations in palm size. In approaches assuming a fixed palm to camera distance, palm area of fixed size is cut out from all the subjects. In such a case, the subjects with palm size smaller than the fixed size are either rejected or are able to infiltrate the system. This problem was raised by Poon *et al.* [139], who proposed the use of variable sized elliptical ROI to address it. By extracting variable sized palm areas, we are able to handle this problem as well.

3.5.3 Image Alignment

Image alignment algorithm was described in Section 3.4. We provide n putative matches as input. The algorithm iterates with a different pair of 4 points everytime. In each iteration, a homography matrix, H_{iter} , is computed with the chosen 4 points pair. An error measure is defined to check the correctness of H_{iter} . Difference of euclidean distance between the actual putative matches and the points transformed using H_{iter} is computed. For each pair of putative match, the pair having an error of less than δ is considered an inlier. The homography resulting in the highest number of inliers is chosen to be correct.

In most cases, we get two images of the same size that are suitably aligned to each other. However, in certain cases, we fail to compute an accurate match. RANSAC is a robust algorithm for Homography computation, and cases where the number of inliers fall beyond a certain number indicate errors in point correspondence. The reason could be that most of the matches had low correlation values. In such cases, we assume that palm textures are different and assign a high dissimilarity value, effectively labeling them as belonging to different individuals.

3.5.4 Feature Extraction

Once the images are aligned, one can use any one of the existing approaches for matching palm-prints. In our experiments, we found the features in frequency domain were more robust to noise, when the frequency was experimentally set at the optimum value. Hence, we match gabor filter responses from corresponding windows and their neighborhood, as proposed in [8]. (For details on gabor filter, please refer to Appendix)

3.5.5 Matching

A typical pose corrected image is shown in Figure 3.4(c). To match the images, we only concentrate on the overlapping parts of the two palms. This is referred to as matching with masking, and the distance is computed as the sum of hamming distances between the binarized real and imaginary filter responses [8]. For certain pairs of genuine matches, error in image alignment can increase the resulting matching score. To overcome this, we define the final dissimilarity score as $d_{final} = \min(d_{fixed}, d_{corrected})$, where the individual dissimilarity scores are computed assuming fixed pose, and after pose correction. If Homography computation is accurate, the dissimilarity score after pose alignment should reduce for genuine pairs.

In short, we compute the best available matching score between any two image pairs. This reduces the variability in genuine matching scores. However, it also improves the matching scores of imposter pairs. The result of this bias towards better match is clearly visible in the results presented in the following section 3.6. The results have been provided in Section. Next, we describe a novel method to address the issue of discarding false line-like impressions.

3.6 Experimental Results and Discussion

We present results of the proposed approach on PolyU dataset with synthetically generated pose variations along with a real dataset with varying palm poses.

3.6.1 Synthetic Dataset

We selected 50 individuals from the PolyU database, and generated various poses. For each subject, we randomly pick 1 image, apply out of plane rotations, and get images from the resulting view. We generated 9 different poses for each subject by rotating the image about X and Y axes (Z being the camera axis), with angles varying from 0° to 45° . Figure 3.6 shows some of the examples from two classes.

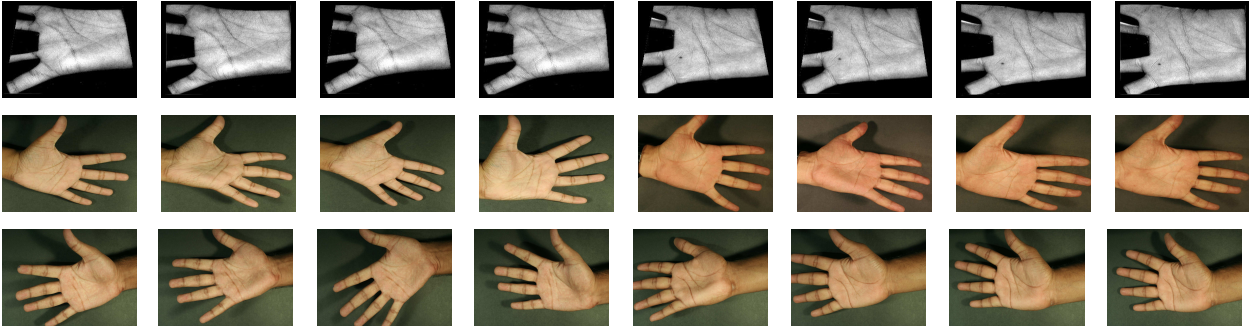


Figure 3.6: A set of images from the Dataset. First row shows 4 images each from 2 users in Synthetic Dataset. 4 left hand images each from 2 users in Real Dataset are shown in the second row and right hand is shown in third row.

To compare our approach with the traditional approach, which assumes a fixed pose (except in-plane rotation and translation), we obtained two values of matching score between every pair, with and without image alignment. Note that the fixed pose approach always extracts a fixed-sized square ROI. Each image was matched with 4 registered images, and the minimum score is taken as the matching score. Comparison of ROC curves between our algorithm and fixed pose approach, for Synthetic Dataset, is given in Figure 3.7(a). A semilog plot of the same is given in figure 3.7(c), to highlight the low FAR regions.

3.6.2 Real Dataset

The synthetic dataset does not model the complexities of unconstrained real world imaging. In reality, variations arising from illumination changes, skin deformation, concave nature of the palm, and specular reflection from the skin, should be considered to check the effectiveness of a matching algorithm(Figure 3.8). We thus created a real world dataset. The setup consists of a fixed camera looking at a uniform background. We collect 5 different palm images with varying poses from each subject. We consider the right and left palms of a user as two different classes, and the dataset consists of 100 classes coming from 50 individuals. However, for the purpose of computation of imposter distances, we do not compare an image from a right palm to a left palm. Figure 3.6 shows four examples each from four classes from the dataset.

For every user, 2 images were used as the registration images. And the rest were used for testing. Two matching scores for each pair were computed using the fixed pose and the proposed approaches. The ROC curves for the two have been presented in Figure 3.7(b) and (d).

Table 3.1 compares the Equal Error Rate while using the two approaches. For synthetic data, we categorized pose variations in 5 groups to analyze performance incrementally. We also present

the results while blindly applying the pose correction using Homography along with the complete matching algorithm. From the results on synthetic data, we note that the pose correction step is inaccurate and introduces additional errors in case of data with small pose variations. However, as the extent of pose variation increases, the correction phase in the proposed approach becomes beneficial, increasing the overall accuracy of the matcher.

Table 3.1: Equal Error Rate(EER) on two datasets with the fixed pose and proposed approach.

Method	Synthetic Data					Real Data
	0°-20°	20°-30°	30°-35°	35°-40°	40°-45°	
Fixed Pose Approach	0.01%	3.24%	3.71%	16.93%	30.92%	22.4%
Blind Pose Correction	16.48%	12.40%	11.14%	14.98%	11.92%	16.51 %
Proposed Approach	0.47%	4.19 %	11.14%	14.98%	11.92%	8.71%

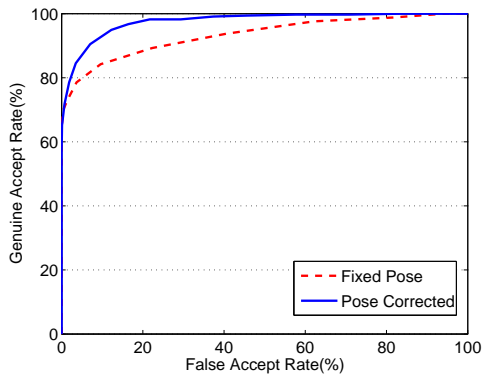
A few interesting characteristics to be noted in ROC plot of real data in Figure 3.7(d):

- (*p*): Note that the GAR drops even with high FAR scores. This indicates the presence of some genuine pairs with low similarity. This can arise in uncontrolled imaging scenarios from variations due to blur, wrinkles, illumination, etc.(Figure 3.8)
- (*q*): The second region of interest is to the left with low GAR, which indicates the presence of image pairs from different palms with very high similarity. This arises mainly due to presence of images with most of the pixels being saturated due to specular reflections from the skin(Figure 3.8(c)).
- (*r*): Note the drop in GAR in pose corrected version is earlier than the uncorrected one. The reason being that pose correction reduces distance between pairs from different palms too, making some of them closer to genuine pair distance.

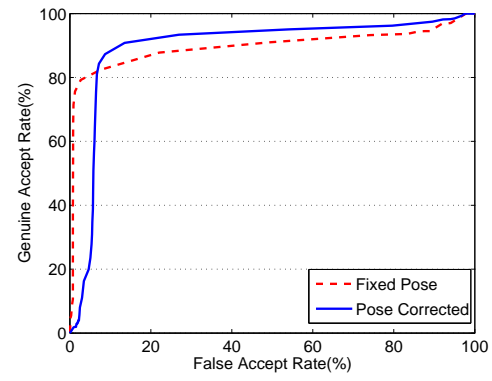
One can rectify some of the above problems by improving the capturing device. For example, to reduce the effects of specular highlights(*p*), use a polarizing filter and improve the gain control of the camera. Improvements in camera sensor and lens can also lead to a reduction in blur(*q*). The last problem is inherent to the approach of pose correction, and requires better matching algorithm to overcome.

3.7 Conclusion and Future Work

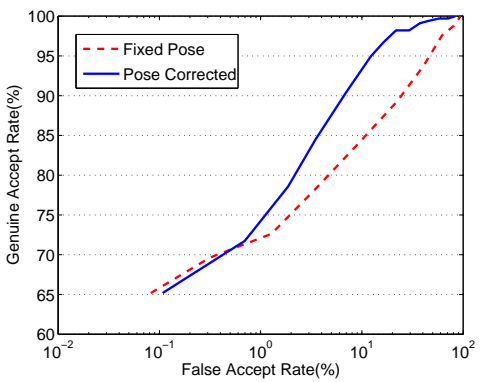
In this paper, we presented a novel approach for enabling unconstrained palmprint recognition over a variety of poses. The dataset we collected contains significant amount of variations due to pose, illumination and specularity. We describe the challenges faced by such a system and proposed a mechanism to address the pose changes. We found that recognition accuracy is better for well-focused images as it is easier to find stable corresponding points across poses. The accuracy of the system is mostly affected by the imposter matches turning good because of a flexible alignment process. Our method achieves considerable reduction in the EER from 22.4% to 8.7% for a set of 100 users.



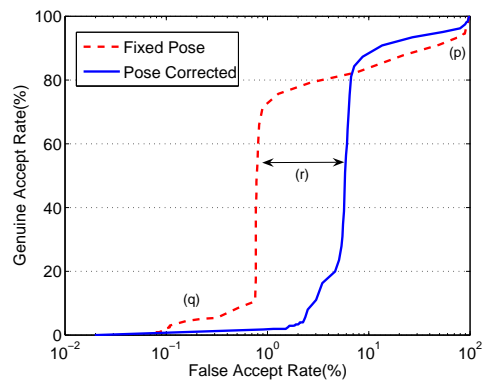
(a) Synthetic Data



(b) Real Data



(c) Synthetic Data (semilog)



(d) Real Data (semilog)

Figure 3.7: ROC curves for the Synthetic and Real Datasets in *linear* and *logarithmic* scales.

One could imagine the use of additional information such as skin color, hand shape, and positions of finger creases, etc. to improve the overall accuracy of matching. One could also improve the accuracy of the approach using better imaging, and using more accurate methods for correspondence computation. One could also employ skin deformation models or using articulated or concave models for the palm surface.



Figure 3.8: Images from our dataset with blur, specular reflection, and skin deformations.

Chapter 4

Video Based Palmprint Recognition

The use of camera as a biometric sensor is desirable due to its ubiquity and low cost, especially for mobile devices. Palmprint is an effective modality in such cases due to its discrimination power, ease of presentation and the scale and size of texture for capture by commodity cameras. However, the unconstrained nature of pose and lighting introduces several challenges in the recognition process. Even minor changes in pose of the palm can induce significant changes in the visibility of the lines. We turn this property to our advantage by capturing a short video, where the natural palm motion induces minor pose variations, providing additional texture information. We propose a method to register multiple frames of the video without requiring correspondence, while being efficient. Experimental results on a set of different 100 palms show that the use of multiple frames reduces the error rate from 12.75% to 4.7%. We also propose a method for detection of poor quality samples due to specularities and motion blur, which further reduces the EER to 1.8%.

4.1 Introduction

As discussed earlier in Chapter 1, with the increasing availability and ease of use of devices having generic web cameras, they are becoming the choicest mediums of image capture for biometric applications to enhance the security in access control for these devices. However, it is extremely challenging to develop algorithms to process low quality images and videos efficiently. When using a generic camera, many factors need to be considered. The illumination is variable, background clutter makes segmentation difficult, the subject may not be aligned properly, or be out of focus. Image quality goes down considerably in such a case. A detailed study on the effects of unconstrained imaging on palmprint recognition has been provided in Chapter 2. We have already provided solutions for some of the problems mentioned here in the previous chapter. Now, we proceed to discuss methods to address illumination variation. Illumination variation and noise caused due to change in pose add extra wrinkles that are interpreted as line like features by the recognition algorithm. We have already talked about the degenerating effects of specular reflections at different angles of the palm in detail in previous sections. It is more pronounced for palm as compared to other biometric modalities like face or gait, since line features get easily distorted with even minor changes in the hand pose(see Figure 4.1).

For robust palmprint recognition in low resolution and partly out of focus images captured in an unconstrained imaging setup, we propose capturing a video instead of a single image. With the ex-

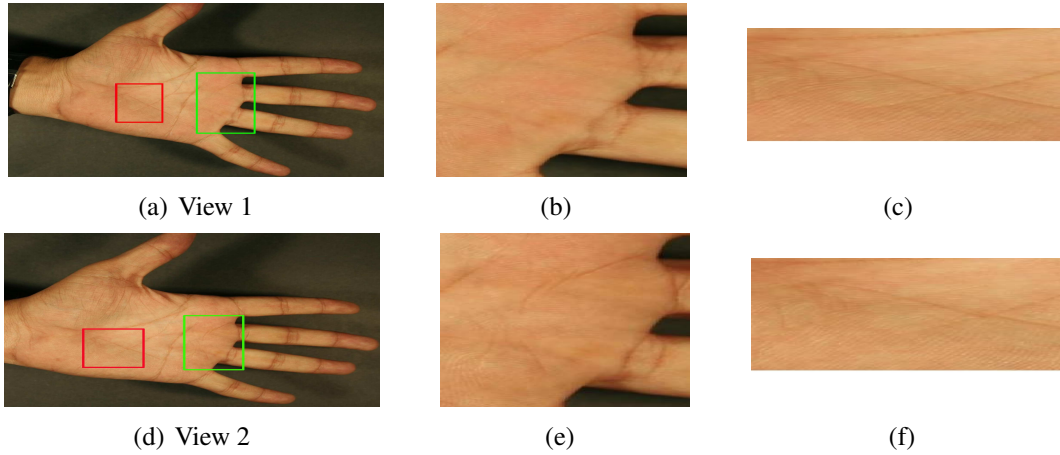


Figure 4.1: Variations in Palm Lines with view variation. The highlighted regions shown separately show that the lines look phased out in different views. c) and e) show clear views.

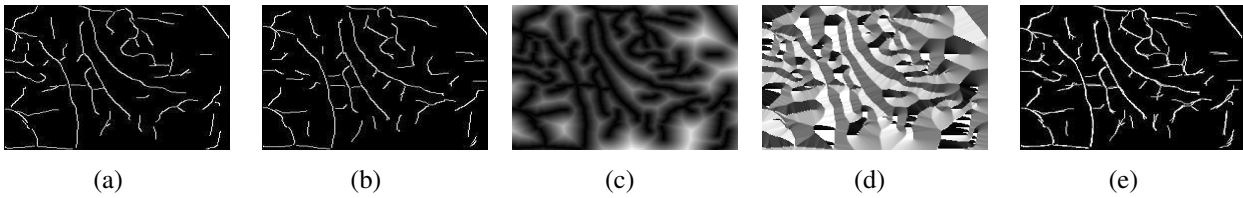


Figure 4.2: The process of Registration: a) Line map of the first image, b) Line map of the second image, c) Euclidean Distance Transform(DT) of the first image, d) Gradient Transform of DT, e) overlapped image

tra information available in the consecutive frames of a video, it is possible to reconstruct an image which retains only the useful information. Most of the consistent lines are identified and the extra wrinkles rejected. There is an added advantage of the system being used for recognition in the large quantities of surveillance data available these days. It is for this reason that video based biometric has recently gained popularity. Various algorithms for video based face recognition [140], and gait recognition [28], have been proposed to achieve view and illumination invariance.

We propose the following framework. At the time of enrollment, a single high quality video of the palm is captured. This is the template palm video. We capture a normal video as the query sequence. We observe that when a user poses the palm for a video, there is an involuntary shaking of the hand. Due to this, certain frames are out of focus, and there is a change in the light reflected from the palm. But even a 1 second video can give us as many as 25 frames. Hence, we can expect, without the loss of generality, that there will be some frames in the video that have well focused palm lines in some part, while there will be other frames that will have clear lines from some other parts of the palm. If this partial information can be combined from all the frames, it can help in reconstructing a single image with stable palm line features.



Figure 4.3: gabor response on successive addition of frames.a) gabor response of the Base Image(BI), b)On adding 2 images to BI, c) On adding 6 images to BI, and d) On adding 10 images to BI

Novelty and Contribution

In this article, we propose, for the first time, a robust palmprint recognition algorithm for recognizing palm that can handle variable illumination settings and noise. The primary contributions of this article include: i) A method to register multiple palm images from a video without relying on correspondences, which are difficult to obtain, ii) A method to integrate the information from multiple frames in the feature space, and iii) A method to detect poor quality acquisitions due to specularities and motion blur so that they can be rejected without comparison. We demonstrate the effectiveness of our approach on a dataset of low quality palmprint videos consisting of a 100 users acquired using a webcam. Our experiments prove that our solution using multiple frames performs better when compared against solution using a single palm image for recognition.

In the next section, we describe the process we use for combining multiple frames and recognizing individuals based on their palms. The rest of the steps can be used directly as described in state-of-the-art algorithms (Figure 3.5). It is assumed here that the users are co-operative, and whatever small projective transformations occur can be assumed to be affine.

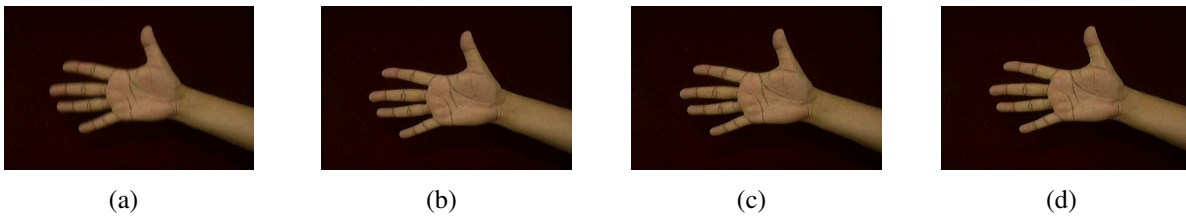


Figure 4.4: Few frames from a video in our dataset showing variations due to illumination and problems in image focusing

4.2 Combination of Multiple Frames

The most common way to reconstruct a single image from multiple images is by using super resolution. A detailed study on super resolution of images has been provided by Farsiu *et al.* in [141]. Registration is a pre requisite for super resolution. The error tolerance in registration step for super resolution should not exceed 1 or 2 pixels. But to achieve such high levels of accuracy in registration, the image should have rich textural information. But this is not possible in the case of

palm images because the only texture present is in the form of weak lines. Also, even if registration is dealt with, the processing time taken to super resolve images is high. This makes it difficult to directly use super resolution for a recognition based application like biometrics.

Arandjelović *et al.* proposed implicit super resolution to achieve pose and illumination invariance in low quality videos [140] for Automatic Face Recognition (AFR). They achieve this by the offline learning of a hierarchy of gSIM models, sub-sampled at multiple scales. In our case, we aim at mainly dealing with missing data. Since the discriminative information present in these images is in the form of lines, it makes sense to reconstruct the image in feature space; where the feature chosen is the line information present in the image. Hence, we propose a novel method to combine the image information in the feature domain. We choose gabor filter response as our working model.

We present a step by step description below:

DATA ENROLLMENT: A high quality video is taken at the time of user enrollment.

MATCHING QUERY VIDEO: The following steps are involved: 1) Frame Extraction from the video: First of all, we need to find all the valid frames from the videos. A valid frame is one which has a clear unobstructed view of the palm. 2) Palm Extraction: The relevant part is extracted from all the frames. This is the first step of registration. This corrects for the in plane rotations, hand orientation is set to a pre fixed direction. 3) Registration: This step employs a registration method which corrects for scale and pose variations. This results in the image being registered within 3-4 pixel range. 4) Combination across multiple frames in the feature space: In this step, we combine the information on a pixel by pixel basis. Registration and frame combination have been described in the Section 4.5. 5) Matching: Since we work with gabor responses, they can directly be used to generate matching scores as described in [8].

4.3 Capturing Palm Video

We used a web camera to obtain the videos. For every palm video, we can extract at most 25 frames per second. Before doing any processing, we need to ensure that the frame gives a clear and unobstructed view of the palm. Our imaging techniques ensures that the palm is in clear view of the camera. The problem regarding the difference in frame capture rate of the camera with the expected rate resulted in the camera buffering the same frame twice. This issue was resolved using a simple background subtraction technique to eliminate the identical image samples.

4.4 Palm Extraction

Palm Extraction is carried out after converting the colour image to a grayscale image. A technique similar to the one followed in Chapter 3 is followed here. The result is the extracted palm from each of the frames of the video.

4.5 Registration and Frame Combination

Registration is the process of overlaying two images taken at different times. Barbará Zitova *et al.* gave a detailed survey of the registration methods currently in use [142]. This alignment is

Algorithm 2 Algorithm for Image Registration

```
1: Input:  $L_1$  and  $L_2$ , line maps of  $Image_1$  and  $Image_2$ , respectively.
2: Output: Transformation matrix  $T_f$ .
3:
4: Initialize:  $iter_{max}$ ,  $dist_{max}$  and set  $T_f$  to identity matrix,
5: Compute  $DT$ , the Euclidean Distance Transform matrix of  $L_1$ .
6: Compute  $f_x$  and  $f_y$ , the Gradient Transform of  $DT$  in both  $X$  and  $Y$  directions.
7: while  $iter < iter_{max}$  do
8:   for all pixels in  $L_2$  do
9:      $[r, c]$  = co-ordinates of the current pixel.
10:     $dist_1 = DT[r, c]$ .
11:    if  $dist_1 > dist_{max}$  then
12:      Point is an outlier.
13:    else
14:      Let  $theta = f_y/f_x$  at  $[r, c]$ .  $[r, c]$  moves to a new position  $[nr, nc]$  such that:
          
$$nr = r + dist_1 \times \sin(theta) ,$$

          
$$nc = c + dist_1 \times \cos(theta) .$$

15:      Append  $[r, c]$  to matrix A and  $[nr, nc]$  to matrix B.
16:    end if
17:  end for
18:   $T_f$  is computed using the following equations given matrices A and B are known:
          
$$A = T \times B ,$$

          
$$T_f = T \times T_f .$$

19:  Transform  $L_2$  using  $T_f$ .
20: end while
21: Output the final transformation matrix  $T_f$ .
```

achieved by following a 4-step procedure: a) Feature detection, b) Feature matching, c) Transform model estimation, and d) Image re-sampling and transformation. Significant regions (forests, lakes), lines (region boundaries, coastlines, roads) or points (region corners, line intersections) are understood as features here. These features must be present in both the images & should be stable with time.

In palmprint images, it is difficult to find robust landmark points since the textural information present in the image is very weak. The local texture around many points on the lines in the image is similar; it is very difficult to say which particular pixel in one image exactly corresponds to which pixel in the other image. Cases of missing and erroneous correspondence computation occur easily; both because there is similarity in the texture and also due to the differences in the computed line map. For such a case, Gut *et al.* in [143] have suggested edge based registration using a Hausdorff distance function modified by using a voting scheme as fitting quality function. Such a method could be used for palm line based registration with some modifications.

In this work, we propose a registration method for palmprints that does not need to find strictly corresponding landmark points. We just need to know the sets of pixels constituting the lines in the two images to be matched. Assuming the underlying transformation to be affine, parameters are estimated iteratively. Loose matches are found using the Euclidean Distance Transform [144]. Outliers are excluded in the process. This method is more flexible as opposed to using strict point correspondences. This implicit adaptivity is extremely helpful in taking account of missing and erroneous data.

We need to register upto 20 frames in every video. Without the loss of generality, we take the

middle frame to be the base image to which the other frames are registered. The algorithm for registration has been provided in Algorithm 2. The details of the line detection method and the method to find point matches has been described below:

Line Detection: In a palmprint image, the major lines are the most natural choice for any alignment process. Thus, the features used for registration are the palm lines. To detect these lines, we employ a sobel filter [137]. The image is filtered in four orientations viz. 0° , 45° , 90° & 135° . A thresholding is then applied to find the peaks found by the filter. This results in the line like structures present in the image like the palm lines and some wrinkles. We reject the wrinkles and keep only the important lines; since small wrinkles are not very reliable features. All the points constituting the lines found in this stage make up the set of pixels to be matched(see Figure 4.2(a) and 4.2(b)).

Finding Correspondences: As we mentioned earlier, finding exact correspondences for a line image is not possible. So we follow a relaxed criterion for correspondence finding. The transformation is assumed to be affine here. Since we find the transformation between the consecutive frames of a video, we can safely assume that the changes in the motion of the hand are continuous in nature. Also, in the frames that are present close to each other, the normal hand motion will be pretty much smooth. Hence, the lines to be matched can be assumed to be in proximity of each other. Thus, we need to find the line pixel closest to any given point in the base image. To find that, we compute the gradient field vector of the euclidean distance transform of the base image. The magnitude of the distance transform gives us the distance and the gradient vector gives us the direction of the closest pixel [143]. So, for every line map pixel in an image to be registered, we can find out the line pixel closest to it in the base image. This is the corresponding point for the pixel in question. Outliers correspond to the wrinkles or lines which do not show up consistently in all the images, and are decided based on a distance threshold. In this way, a loose set of correspondences for every line is found.

Model Estimation and Image Transformation: After the loose correspondences are found, a simple pseudo inverse based technique is applied to find out the affine transformation parameters. The image is transformed for the next iteration and the same process continues till the stopping condition is reached. We finally get a pair of registered images with an average pixel error of 3-4 pixels(see Figure 4.2(e)).

Frame Combination: Once the images have been registered, we need to combine them in a meaningful manner. The basic idea behind the combination scheme is that the consistent information that keeps showing up in all the images must be preserved, and all the elements of noise will be such that they will appear and disappear through the various frames. This happens due to the small involuntary movements made by the hand. Because of the change in the specular reflection from the surface of the skin, different parts seem to be in focus in different frames. Hence, to retain the relevant information, we keep on adding the gabor responses at every pixel. This results in a kind of weighted average, such that, the sequence being repeated the most gets the highest weight. This results in the major lines appearing in all the frames getting high weights, as compared to the scarce wrinkles that are almost negligible. It is after this weighted averaging that we threshold the image, as a result of which pixels with the least weights are shedded away(see Figure 4.3). This image is then used for the final matching score computation. This method of frame combination has been chosen as it gives the best results as compared to other strategies (See the results section 4.6).

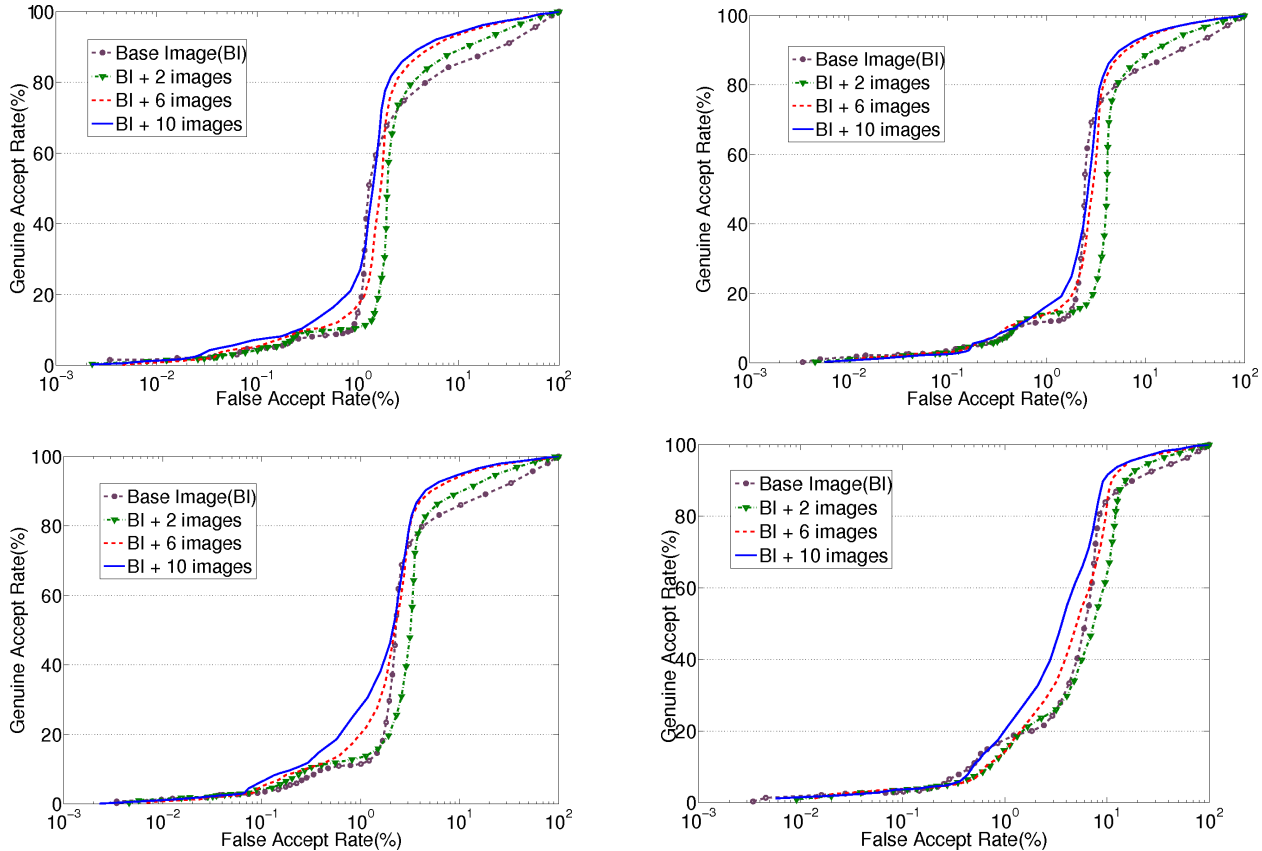


Figure 4.5: ROC curves on *logarithmic* scale for a) *Parameter Set*₁, b) *Parameter Set*₂, c) *Parameter Set*₃, d) *Parameter Set*₄.

4.6 Dataset, Experimental Results and Discussions

The Dataset contains videos of the hand taken by a commonly available web camera, in an unconstrained image capturing setup. This is the first video database containing palmprint videos. 100 subjects were asked to pose for a fixed camera in a manner intuitive to them. 6 videos each were recorded for both the left and the right hand for each subject. Few frames from a video in our database have been shown in Figure 4.4.

Experiments: We conducted a variety of experiments on samples from the dataset described above. We performed experiments on a set of 100 subjects. The matching between two palms, $palm_1$ and $palm_2$ is performed by computing the two hamming distances R_m and I_m between the real and imaginary gabor responses (See Appendix). The final distance score is taken to be the maximum of these two dissimilarity scores. We conducted the following sets of experiments:

Experiment 1: Firstly, we wanted to find out the set of optimal gabor filter parameters suited for the current experiment. Please note that the gabor filter parameters need to be attuned to the database to find out the best performance. We used 4 sets of different gabor filter parameters to obtain the best result. The 4 parameters needed to fully determine a gabor filter are *window size*(w), *Gaussian variance*(σ), *frequency of the sine wave*(u) and *orientation angle of the filter*(θ). (See Appendix) For all the parameter

Table 4.1: Equal Error Rate(EER) observed using different gabor filter parameters.

	Base Image	On adding 2 frames	On adding 4 frames	On adding 6 frames	On adding 8 frames	On adding 10 frames
<i>Parameter Set₁</i>	15.47%	12.92 %	7.39%	7.40%	8.74%	6.01%
<i>Parameter Set₂</i>	12.92%	9.87 %	7.32%	8.90%	7.17%	7.47%
<i>Parameter Set₃</i>	10.26%	8.68 %	7.94%	7.58%	8.29%	6.22%
<i>Parameter Set₄</i>	12.28%	13.41 %	11.67%	10.61%	9.64%	9.13%

sets, θ was set to the following values $[0^\circ, 45^\circ, 90^\circ, 135^\circ]$. The maximum response over θ at every pixel was recorded. The other parameters used are these: 1) $w = 27, \sigma = 6.4, u = 0.08$, 2) $w = 29, \sigma = 6.4, u = 0.075$, 3) $w = 31, \sigma = 7.1, u = 0.07$, 4) $w = 35, \sigma = 7.7, u = 0.06$.

In this experiment, we combine the images by first arranging all the samples in increasing order of each pixel value. Then, we remove the maximum and minimum value as the extreme cases are generally considered to be noisy. Finally, we take the mean of the remaining values. A total of 3528 genuine match scores and 175065 imposter match scores were recorded for each parameter set. Table 4.1 shows the results obtained for each parameter set on a set of 50 subjects using both their right and left hand images. Generally, both right and left palm images are considered to be different users in palmprint recognition. Hence, we effectively record results for 100 users treating the right and the left hand templates as different users. We show the effect of incrementally adding frames to a base image. We provide results for reconstruction done by adding frames upto a maximum of 11. In each of the cases, we observe that the final result after adding all the 11 frames is better than using only the base image. This proves our claim that if the frames of a video are properly combined, the accuracy obtained is better than using only 1 image. The related ROC plots on a linear and semilog scale are shown in Figure 3.7. Please note that the semilog scale is used to provide emphasis on the lower FAR region where the difference in accuracies can generally be seen. *We choose the Parameter Set 1 for further experiments.*

Table 4.2: EER and FTA (in %) with varying number of frames and different combination strategies. The results have been computed for 30 users.

	Base Image (BI)		BI+2 frames		BI+6 frames		BI+8 frames		BI+10 frames	
	EER	FTA	EER	FTA	EER	FTA	EER	FTA	EER	FTA
M_1	18.24	1.8	16.05	0.93	19.34	1.40	23.53	1.40	26.86	1.40
M_2	18.24	1.8	15.03	0.93	15.48	1.40	19.15	1.4	19.39	1.40
M_3	18.24	1.8	10.94	1.4	9.55	0.9	10.56	0.93	9.50	0.93

However, in Figure 3.7, in all the curves, we observe a slow rise in GAR initially. This happens due to the texture appearance being partially washed out due to specularity. In Table 4.1, we observe that the EER does not reduce monotonically with the increase in number of frames being added. This could be due to the fact that some noise also keeps getting added as we add more and

more frames to the Base Image as well as the blurring of lines arising from misalignments during the registration phase. These are some of the challenges associated with uncontrolled imaging setups. The problem can partially be solved by using sophisticated combination strategies, but those end up increasing the computation time of the algorithm. The current algorithm takes 0.8 seconds to reconstruct an image using 11 frames. This effect can however be eliminated by automatically detecting and discarding these washed out samples at the time of query itself. This is termed as the Failure To Acquire(FTA) rate [1.4.2](#).

Experiment 2: We performed another experiment to study the effect of variation of the FTA rate on the performance of the recognition system. We determine the washed out samples by measuring the average response of the image to gabor filter. If the response lies below a particular threshold, we say that the image does not have the requisite texture information. Using the best set of gabor filter parameters found in the previous experiment, and a cutoff threshold on top of it, we vary the frame combination strategy to figure out the best combination that provides the best result. The FTA rate is different at various levels of frame addition on account of rejecting samples at each stage independently.

For frame combination, at every given pixel, we combine the information by taking all the values at that exact same pixel location in all the x frames being considered upto that point(x lies between 1 and 11), and then combine these values using a combination strategy. We use the following combination methods: i) M_1 : We consider the maximum of all the values corresponding to a given pixel. This is done in expectation of the fact that the gabor response will maximize at the occurrence of a line. ii) M_2 : In general, in most data combination strategies, to find the group behavior, the extreme values(i.e maximum or minimum) are considered as noise and only the remaining values are considered. In our case, since small noise patterns can cause unwanted peaks in the data, we compute results using the second maximum value out of all the corresponding values at a point. iii) M_3 : We take the mean of all the values present at a particular pixel to model the average behavior of the curve at that point. These strategies were applied on a set of 30 users and the results obtained are shown in [Table 4.2](#). The table shows the results of strategy vs. number of frames combined to produce the result.

Here, we can clearly see that after using 11 frames, taking the average of all the frames is able to represent the data better than the maximum or the second maximum values. The results using the average of all 11 frames is 9.50 as opposed to the result using maximum value, 26.86, or that obtained using the second maximum value of 19.39. Hence, we combine the frames using an average of frames in the experiments to follow. It can also be seen, by comparing the first rows of [Table 4.1](#) and [Table 4.3](#) that taking the average of all the frames provides better results than the combination method used in Experiment 1.

Experiment 3: Our next experiment aims to observe the effect of the thresholds used on the Equal Error Rate. Hence, we use three different sets of thresholds to obtain results using the average combination strategy with the gabor filter parameters found in Experiment 1. We used three different thresholds to study the effect of removing bad samples from the dataset incrementally. These three different parameter sets namely, τ_1 , τ_2 and τ_3 put a cutoff limit on the mean and variance of the gabor filter responses of the palm images. The idea is that the washed out images shall give a lower gabor filter output as well due to a lack of presence of texture information. Thus, we put thresholds on the first and second order moments (namely the mean and variance) of the output to weed out images with very low texture output. The thresholds τ_1 , τ_2 and τ_3 are in increasing order of the cutoffs, respectively. The results for these three different parameters namely, τ_1 , τ_2

and τ_3 have been provided in the Table 4.3. The second row in the table shows the results on not applying any cutoffs. Please note that the FTA rate is different at various levels of frame addition on account of rejecting samples at each stage independently.

Table 4.3: EER and FTA (in %) with varying number of frames and different quality thresholds.

	Base Image (BI)		BI+2 frames		BI+6 frames		BI+8 frames		BI+10 frames	
	EER	FTA	EER	FTA	EER	FTA	EER	FTA	EER	FTA
SF	12.75	0	24.30	0	19.25	0	17.48	0	36.19	0
τ_0	12.75	0	13.99	0	4.70	0	5.15	0	5.79	0
τ_1	7.64	6.57	8.07	8.09	5.35	7.58	4.42	7.5	4.50	8.09
τ_2	14.36	9.44	5.82	13.15	4.64	10.9	4.46	11.12	3.62	11.8
τ_3	3.69	11.8	3.40	16.86	1.90	13.99	1.80	13.99	7.75	14.5

The first row in Table 4.3 shows the result for fusion of scores(SF) obtained by matching frames directly without combination. The max rule was used to obtain these results. The frames were matched directly and the maximum of the scores was taken. The second row of Table 4.3 shows the improvement in the EER(Equal Error Rate) as we add more frames from the video to the Base Image. The corresponding ROC curve has been shown in Figure 4.6. We note that the EER first drops from 12.75% to 4.7% with the addition of 6 frames to the base image. However, the addition of further frames decreases the accuracy. This could be due to the noise present in the frames, as well as the blurring of lines arising from misalignments during the registration phase.

On observing the curve in semilog axis(4.6(b)), we notice a slow rise in the GAR initially. A similar behavior of the curve was observed in 3.6. The drop indicates the presence of few imposter scores having better matching scores than genuine scores. This happens due to the texture appearance being partially washed out due to specularly. This is characteristic of an unconstrained imaging system. This effect can however be eliminated by automatically detecting and discarding these washed out samples at the time of query itself. This is termed as the Failure To Acquire(FTA) rate. Our final experiment consists of studying the effect of variation of the FTA rate on the performance of the recognition system. Figure 4.6(c) shows the resulting ROC curve, where the EER drops to 1.8% on combining 9 frames. We note that the imposters with high matching score have been completely removed in the process.

To summarize the experiments section, we performed experiments on a set of 100 different palms having an average of 6 videos each. We performed three experiments. The first experiment determines the set of gabor parameters optimal for our database. The second experiment finds out the best combination strategy applicable to us. In the third experiment, we use a set of threshold cutoffs to obtain the Failure to Acquire rate, which automatically detects samples that are beyond the scope of image improvement algorithm used by the system and alerts user to provide a different sample. The proposed algorithm takes only 1.4 seconds to combine 11 frames for a MATLAB implementation as compared to a few minutes taken for the fastest and most optimized super resolution techniques for these images. A total of 3, 528 genuine match scores and 1, 75, 065 imposter match scores were recorded.

4.7 Conclusion

We address the problem of palmprint recognition from low-resolution videos of the palm captured in an unconstrained setup using low-end cameras. We proposed efficient and robust methods for feature level integration of information from multiple frames. We show that the EER of an authentication system reduces from 12.75% to 4.7% by integrating information from just 9 frames, on a dataset of 600 samples from 100 palms. We also propose a method to detect and remove low quality captures, where the texture information is washed out, which further enhances the EER to 1.8% albeit with rejection of some samples. The resulting system is both efficient and robust to be practical.

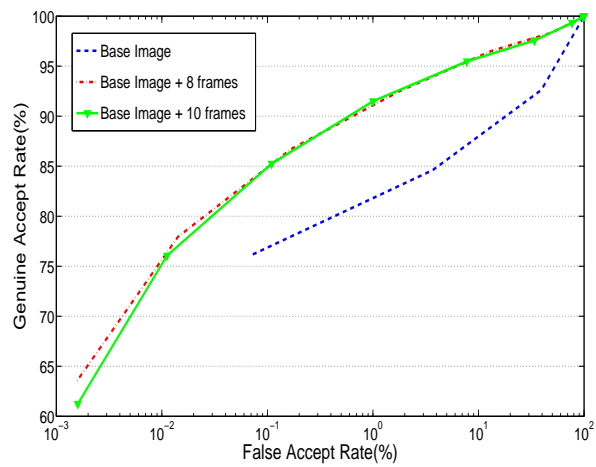
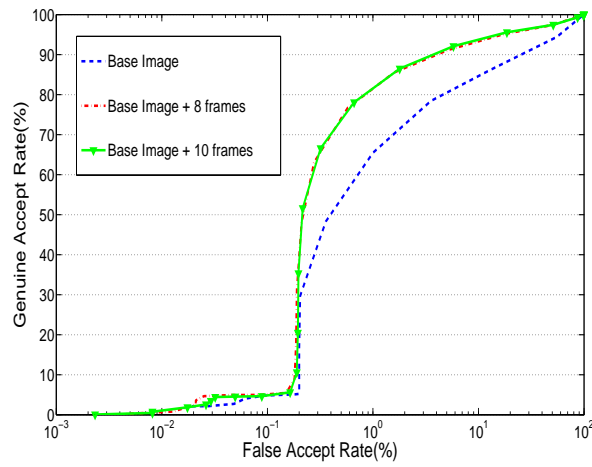
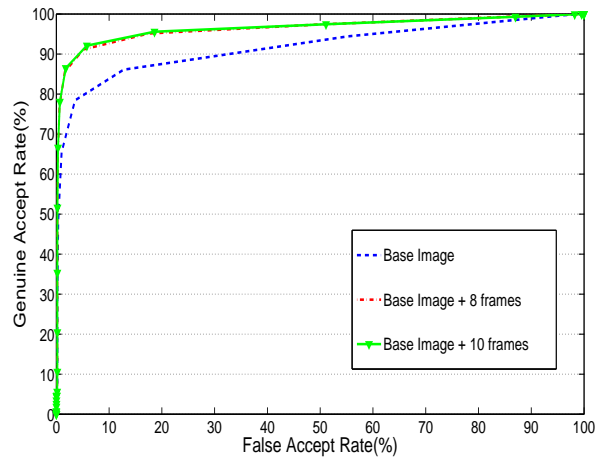


Figure 4.6: ROC curves on a) *linear* and b) *logarithmic* scales, and the result of removing poor captures.

Chapter 5

Conclusions and Future Work

The potential of access control using biometrics in the field of civilian applications is huge and largely unexplored so far. The prime reason for this lag is the challenges faced while handling biometric templates captured using unconstrained and unsupervised camera based imagery associated with civilian usage.

We try to solve some of the challenges faced in developing biometric applications for civilian use. We envision that the most practical image capturing module for a civilian application will be the generic web cameras embedded with the laptops and mobile cameras available these days. These cameras are low in quality as compared to the traditional image capturing devices. The information contained in these low quality images is further distorted by the presence of illumination and contrast variations due to surrounding and also due to the pose variations induced by the improper template presentation by the user. These image degradations prove to be critical for the process of person authentication since even small details in appearance contribute to identifying an individual uniquely.

In this thesis, we have proposed novel algorithms to identify individuals using palmprints that are invariant to pose, illumination and contrast changes. Pose invariance has been achieved by transforming the test template to the template stored in the database at the time of the enrollment. For this, palm is assumed to be a planar surface which reduces the transformation matrix to a linear equation. To make the process illumination invariant, we captured short videos of palm instead of single images and used the information present in the consecutive frames to single out real lines from the artificial lines created by shadows, wrinkles, loose skin etc.

Pose Invariant Palmprint Recognition: To validate our results for pose invariant algorithm, we created a synthetic dataset having artificially induced pose variations. This allowed us to study the effect of pose variations independent of the illumination and contrast variations. To study the effect of real world conditions and to validate experiments on real data, we captured two separate datasets; one having both pose and illumination variations, and another having illumination and contrast variations in low quality images. Each dataset has a total of 100 users; which is a reasonable number for validating biometric experiments. The first dataset has 5 images for each user having a variation of pose angles ranging from 0° to 45° with respect to both X and Y axes. (Here 0° refers to the plane parallel to the camera imaging plane). The users were asked to pose intuitively in different poses. This dataset captures the real-life variations well including some extreme cases in the range of 30° and more. Our experiments prove the efficacy of our algorithm as we are able to reduce the Equal Error Rate on synthetic dataset from 30.92% to 11.92% and on real

dataset 22.08% to 8.71%.

One could improve our approach for pose invariant palmprint recognition by not generalizing the palm surface as a plane, but model as a complex equation of 3D surface/mesh. It could also be divided in two parts with one part being approximated with a plane, and the other region near the thumb could be represented with a concave model. It is also possible to capture a higher number of templates for each user at the time of enrollment and learn through these poses and be able to extrapolate the missing poses and hence populating the pose matrix in a certain expected range. At the time of testing, the test template can just be matched to the entire pose matrix and the minimum score obtained can be assigned as the matching score of the two different poses. There is scope to further improve the point matching technique.

Illumination and Contrast invariance for Palmprint Matching: The dataset containing low quality palm videos of palms having illumination and contrast variation has been captured for 100 users. This dataset allows us to study the impact of all factors on palmprint recognition except pose variation. This is important as a user can be taught to present the palm in the right pose but the illumination variations around him/her cannot be controlled. Our algorithm achieves a reduction in Equal Error Rate from 12.47% to 4.7%. The algorithm can also automatically detect if the video sample is too bad to be used. In such a case, the user can be alerted to provide a better sample. Thus, giving us a final Equal Error Rate of 1.8%, which can be compared to the state of art solutions using restrictive imaging systems. Hence, our algorithm is extremely robust to the effect of surrounding, and ensures that the test template is either improvable, in which case it uses the improved/modified template, otherwise it just rejects the template.

One could improve the registration accuracy and the image combination process. A detailed study needs to be done on the effect of adding more frames for combination. This method could also be combined with the pose correction approach developed in the previous algorithm to make a robust system with further improved accuracy.

Appendix

5.1 Gabor Filter

Gabor filter is a linear filter which acts as a band pass filter for signals. It can be used for edge detection in images. Both, 1D and 2D Gabor filters are Gaussian functions modulated by sinusoidal waves. In both the cases, a mother wave can be used to generate self-similar waves by changing its orientation or spatial shift.

5.1.1 1D Gabor filter

The equation of a 1D Gabor filter(See Figure 5.1) can be given as:

$$g(t) = ke^{j\theta}w(at)s(t) \quad (5.1)$$

where,

$$\begin{aligned} w(t) &= e^{-\pi t^2}, \\ s(t) &= e^{j(2\pi f_o t + \theta)}, \\ e^{j\theta} s(t) e^{j(2\pi) f_o t + \theta} &= (\sin(2\pi f_o t + \theta), \cos(2\pi f_o t + \theta)) \end{aligned}$$

Here, k , θ , f_o are filter parameters. Gabor filter is a complex filter, which means that it has two out of phase parts, one is real and another imaginary. The real part holds the following filter:

$$g_r(t) = w(t)\sin(2\pi f_o t + \theta) \quad (5.2)$$

and the imaginary part has the following filter:

$$g_i(t) = w(t)\cos(2\pi f_o t + \theta) \quad (5.3)$$

5.1.2 2D Gabor filter

The 2D Gabor filter, which is what is used for the purpose of edge/line segmentation in images is presented below. It can be represented as the multiplication of the complex 2D sinusoidal wave and a 2D Gaussian kernel (See Figure 5.2 [145]).

$$g(x, y) = s(x, y)w_r(x, y) \quad (5.4)$$

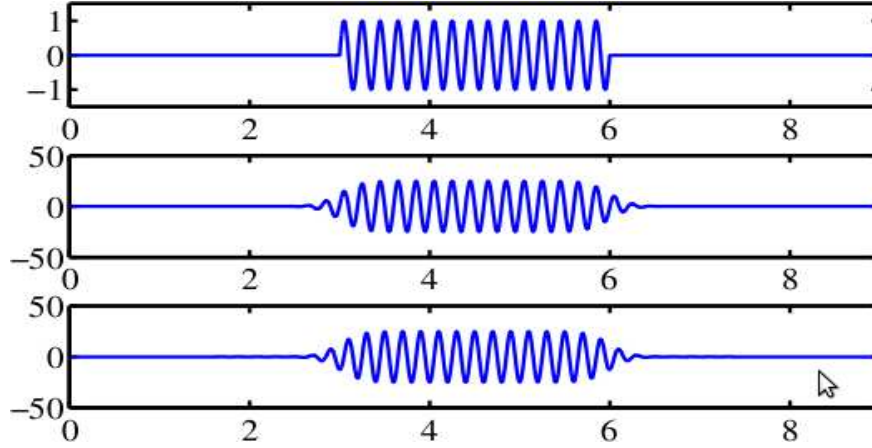


Figure 5.1: a) an input signal, b) output of Gabor filter(cosine carrier) and c) output of Gabor filter(sine carrier)

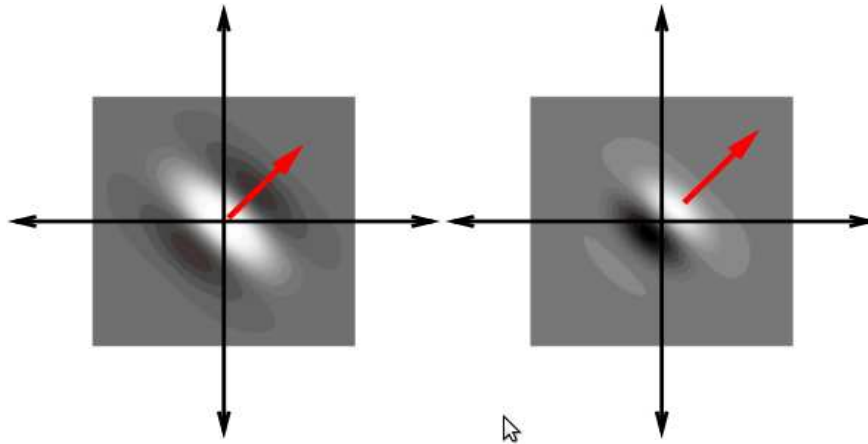


Figure 5.2: The real and imaginary parts of a Gabor filter.

where $s(x, y)$ is a complex sinusoid, also known as the carrier and $w_r(x, y)$ is a 2D Gaussian-shaped function, also called the envelope.

The complex sinusoid carrier The equation of the 2D complex sinusoid carrier can be gives as follows:

$$s(x, y) = \exp(j(2\pi(u_0x + v_0y) + P)) \quad (5.5)$$

where (u_0, v_0) and P define the frequency and the phase of the sinusoid respectively. The sinusoid can be thought of as a double sinusoid one being the real and the other being the imagery part with the following equations.

$$\text{Re}(s(x, y)) = \cos(2\pi(u_0x + v_0y) + P), \text{Im}(s(x, y)) = \sin(2\pi(u_0x + v_0y) + P) \quad (5.6)$$

The Gaussian Envelope The equation of the Gaussian envelope is as follows:

$$w_r(x, y) = K \exp(-\pi(a^2(x - x_0)_r^2 + b^2(y - y_0)_r^2)) \sin\theta, \quad (5.7)$$

where (x_0, y_0) is the peak of the function, a and b are scaling parameters of the Gaussian and the $_r$ subscript stand for an orientation rotation parameter.This can be given as:

$$(x - x_0)_r = (x - x_0) \cos\theta + (y - y_0) \sin\theta, (y - y_0)_r = (x - x_0) \sin\theta + (y - y_0) \cos\theta \quad (5.8)$$

A summary of all the parameters of complex gabor filter can be given as follows:

- **K** : Scales the magnitude of the gaussian envelope.
- θ : Rotates the angle of orientation of the filter. This can be used to detect edges at different angles, like, perpendicular, horizontal etc.
- (a, b) : Scales the X- and Y-axes of the envelope.
- (x_0, y_0) : Location of the peak of the gaussian curve.
- **P**: Phase of the sinusoidal carrier wave.
- (u_0, v_0) : Spatial frequency of the sinusoidal wave.

5.2 RANSAC Based Homography

There are two different concepts involved when we talk about RANSAC based homography. One is the transformation that gives us the mapping between two planes, called homography. The other is RANSAC, which is an algorithm that fits any given model given a set of training points in the case when the training data is noisy. Both of these concepts have been defined in more detail in the following section.

5.2.1 Homography

When a planar object is viewed from two different cameras, C_1 and C_2 , then any point P on the plane gets projected as P_1 when viewed by C_1 , and as P_2 when viewed by C_2 . Here, we consider the camera C_1 to be at the origin and the camera C_2 has a translation of t and a rotation of R relative to the origin or C_1 (See Figure 5.3 [135]). Let the 3D cartesian co-ordinates of the point on the plane be P :

$$P = (X, Y, Z) \quad (5.9)$$

The desired transformation from C_1 to C_2 can be given as:

$$H = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \quad (5.10)$$

where R and t are the rotation matrix and the translation vector respectively. The resulting point is $P' = [X'Y'Z']$ is transformed a homogenous point. Hence, the transformation between the homogenous points P_1 and P_2 can be given by:

$$P' = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} P \quad (5.11)$$

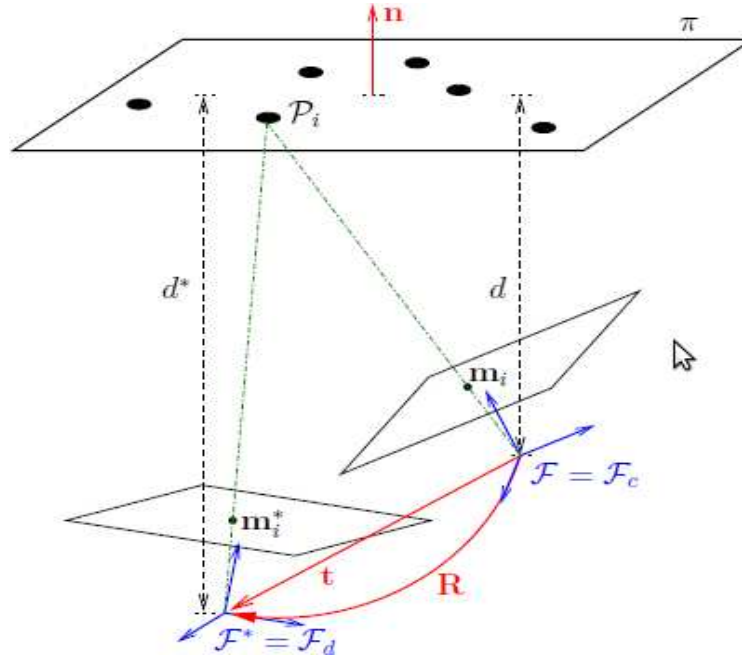


Figure 5.3: Planar points as viewed by two different cameras

The above matrix has 8 parameters to scale (this means that the parameters can be reduced from 9 to 8 by dividing all terms by h_{33} , giving a scaled value of the result). The equation formed can be solved by four point correspondences.

5.2.2 RANSAC

The RANSAC (RANdom Sample And Consensus) is a method to estimate the parameters of a certain model starting from a set of data contaminated by outliers. It is an iterative procedure in which the “true” model is first initialized. The outliers are then estimated based on an error threshold using some criterion function on the actual model. The model function is then refined based on the obtained inliers. This iterative improvement is continued till some pre defined threshold. The percentage of outliers which can be handled by RANSAC can be larger than 50% of the entire dataset. Such a percentage, known also as the breakdown point, is commonly assumed to be the practical limit for many other commonly used techniques for parameter estimation (such as the least squares or median of squares etc.) ([146]).

In our case, the estimated model is that of planar homography. We require only 4 accurate point correspondences for determining the homography. But, what we can actually get from the image is a set of $n > 4$ approximate point correspondences. Hence, we use a RANSAC based approach on this dataset to weed out the outliers and find the approximate underlying model. This is done by an iterative two step procedure as defined below.

- **Hypothesize.** First, minimal sample sets (MSSs) are randomly selected from the input data and the model parameters are computed using only the elements of this MSS. The MSS here, consists of 4 points. A homography matrix is first initialized by randomly selecting four points. This is our “true” model.

- **Test.** In the second step, RANSAC checks which elements of the entire dataset are consistent with the model instantiated with the parameters estimated in the first step. The data points which are found to be consistent with the model are then used for randomly sampling the next 4 points for model computation. It is also called the Consensus set.

The algorithm terminates when a pre-defined number of inliers have been found.

Related Publications

- Chhaya Methani, Anoop M. Namboodiri, “Pose Invariant Palmprint Recognition”, in *Proceedings of the Third IEEE International Conference on Biometrics (ICB)*, 2009, pp. 577-586.
- Chhaya Methani, Anoop M. Namboodiri, “Video Based Palmprint Recognition”, to appear in *IEEE/IAPR International Conference on Pattern Recognition (ICPR)*, 2010.

Bibliography

- [1] http://www.cilab.upf.edu/biosecure1/public_docs_deli/BioSecure_Deliverable_D10-2-3_b3.pdf. 3
- [2] <http://www.internetworldstats.com/stats.htm>. 3
- [3] <http://www.davesite.com/webstation/net-history.shtml>. 3
- [4] <http://en.wikipedia.org/wiki/Biometrics>. 4
- [5] <http://www.tns.com/biometrics.asp>. 4
- [6] N. K. Ratha, J. H. Connell, and R. M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems,” *IBM systems Journal*, vol. 40, pp. 614–634, 2001. 4
- [7] www.biometrics.gov/Documents/biointro.pdf. 4, 7
- [8] D. Zhang, W.-K. Kong, J. You, and M. Wong, “Online palmprint identification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1041–1050, 2003. 7, 21, 26, 29, 31, 32, 34, 46, 48, 49, 55, 56, 57, 65
- [9] J. A. Swets, *Signal detection theory and ROC analysis in psychology and diagnostics : collected papers*. 1996. 11
- [10] http://en.wikipedia.org/wiki/Receiver_operating_characteristic. 11
- [11] J. hanley and B. J. McNeil, “A method of comparing the areas under receiver operating characteristic curves derived from the same cases,” *Radiology*, vol. 148, no. 3. 11, 13
- [12] J. Fogarty, R. baker, and S. Hudson, “Case studies in the use of roc curve analysis for sensor-based estimates in human computer interaction,” 13
- [13] McClish and D. Katzman, “Analyzing a portion of the roc curve,” *Medical Decision Making*, vol. 9, no. 3, pp. 190–195, 1989. 13
- [14] Dodd and E. Lo, “Partial auc estimation and regression,” *Biometrics*, vol. 59, no. 3, pp. 614–623, 2007. 13
- [15] <http://users.soe.ucsc.edu/~karplus/papers/better-than-chance-jul-07.pdf>. 13
- [16] Y. Zhu, S. C. Dass, and A. K. Jain, *Compound Stochastic Models For Fingerprint Individuality*. 2006. 15

- [17] B. Y. Hiew, A. B. J. Teoh, and D. C. L. Ngo, *Automatic Digital Camera Based Fingerprint Image Preprocessing*. 2006. 16
- [18] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, 1991. 16, 17, 25, 37, 48
- [19] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997. 16
- [20] S. Li and A. K. J. (Eds.), *Handbook of Face Recognition*. 2005. 17
- [21] http://en.wikipedia.org/wiki/Facial_recognition_system. 17
- [22] <http://www.technologyreview.com/Infotech/18796/?a=f>. 17
- [23] http://en.wikipedia.org/wiki/Iris_recognition. 17
- [24] J. Daugman, *How iris recognition works*, vol. 14. 2004. 17
- [25] http://en.wikipedia.org/wiki/Retina_scan. 17
- [26] T. G. Group, *Retina and Iris Scans*. 17
- [27] <http://www.gait.ecs.soton.ac.uk/>. 18
- [28] A. Kale and A. R. Chowdhury, "Towards a view invariant gait recognition algorithm," *IEEE Conference on AVSS*, pp. 143–150, 2003. 18, 25, 37, 48, 63
- [29] http://en.wikipedia.org/wiki/Keystroke_dynamics. 18
- [30] M. D. Shanmugapriya and D. G. Padmavathi, "A survey of biometric keystroke dynamics: Approaches, security and challenges," *International Journal of Computer Science and Network Security*, vol. 5, no. 1, 2009. 18
- [31] C. V. Jawahar, A. Balasubramanian, M. Meshesha, and A. M. Namboodiri, *Retrieval of online handwriting by synthesis and matching*, vol. 42. 2009. 19
- [32] http://en.wikipedia.org/wiki/Speaker_recognition. 20
- [33] S. E. Tranter and D. A. Reynolds, *An overview of automatic speaker diarization systems*, vol. 14. 2006. 20
- [34] A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain, "Personal verification using palmprint and hand geometry biometric," *AVBPA*, pp. 668–678, 2003. 20, 31
- [35] http://www.ncstl.org/evident/Jan09_Mitchell_ResearchFocus. 21
- [36] M. G. K. Ong, C. Tee, and A. T. B. Jin, "Touchless palmprint biometric system," *International Conference On Computer Vision Theory and Applications*, vol. 2, 2008. 25, 37, 38, 39, 49

- [37] Y. Han, Z. Sun, F. Wang, and T. Tan, "Palmprint recognition under unconstrained scenes," *ACCV*, vol. 2, pp. 1–11, 2007. [25](#), [37](#), [39](#), [40](#), [50](#)
- [38] A. Kumar, "Incorporating cohort information for reliable palmprint recognition," *ICVGIP*, pp. 583–590, 2008. [25](#), [37](#), [39](#), [50](#)
- [39] N. Duta, A. K. Jain, and K. V. Mardia, "Matching of palmprints," *Pattern Recognition Letters*, vol. 23, pp. 477–485, 2002. [26](#), [29](#), [49](#)
- [40] <http://www.biometrics.gov/Documents/PalmPrintRec.pdf>. [28](#)
- [41] Z. Tan, J. Yang, Z. Shang, G. Shi, and S. Chang, *Minutiae-Based Offline Palmprint Identification System*. 2009. [28](#), [48](#)
- [42] <http://www.horiba.com/scientific/products/forensics/afisapis/faqs/>. [28](#)
- [43] <http://www.silicon.com/management/public-sector/2006/03/23/police-get-national-biometric-palm-print-database-39157511/>. [28](#)
- [44] D. Zhang and W. Shu, "Two novel characteristics in palmprint verification: datum point invariance and line feature matching," *Pattern Recognition Letters*, vol. 32, no. 4, pp. 691–702, 1999. [29](#)
- [45] <http://www.printrakinternational.com/omnitrak.htm>—*Printrak Automatic Palmprint Identification System*, 2003. [29](#)
- [46] D. Zhang, *Automated Biometrics-Technologies and Systems*. Boston:Kluwer Academic, 2000. [29](#)
- [47] C. C. han, "A hand-based personal authentication using a course-to-fine strategy," *Pattern Recognition Letters*, vol. 22, no. 11, pp. 909–918, 2004. [29](#), [31](#), [49](#)
- [48] <http://www.cse.unr.edu/~bebis/CS791E/Notes/Thresholding.pdf>. [31](#)
- [49] X. Pan and Q. qi Ruan, "A modified preprocessing method for palmprint recognition," *ICSP Proceedings*, 2006. [31](#)
- [50] Y. Wang, Q. Ruan, and X. Pan, "An improved square-based palmprint segmentation method," *Proceedings of 2007 International Symposium on Intelligent Signal Processing and Communication Systems*, 2007. [32](#), [34](#)
- [51] W. X. Li, "Authenticating personal identities using palmprint recognition," *PhD Thesis, Hong Kong Polytechnic University*, 2004. [32](#)
- [52] C. C. han, H. Cheng, and K. Fan, "Personal authentication using palmprint features," *Pattern Recognition Letters*, vol. 36, pp. 371–381, 2003. [32](#)
- [53] R. S. Chora and M. Chora, "Hand shape geometry and palmprint features for the personal identification," *Proceedings of the International Conference on Intelligent Systems Design and Applications*, vol. 6, 2006. [33](#)

- [54] C. Poon, D. Wong, and H. Shen, "A new method in locating and segmenting palmprint into region-of-interest," *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. [34](#), [35](#)
- [55] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognition Letters*, vol. 42, pp. 1408–1418, 2009. [34](#), [43](#), [45](#)
- [56] P. A. R. Rodrigues and J. D. L. Silva, "Biometric identification by dermatoglyphics," *Proceedings of Intl. Conference on Image Processing*, vol. 1, pp. 319–322, 1996. [34](#)
- [57] S. Y. Kung, S. H. Lin, and M. Fang, "A neural network approach to face/palm recognition," *Proceedings of IEEE Workshop on Neural Networks for Signal Processing*, pp. 323–332, 1995. [35](#)
- [58] D. S. Huang, W. Jia, and D. Zhang, "Palmprint verification based on principal lines," *Pattern Recognition*, vol. 4, pp. 1316–1328, 2008. [35](#)
- [59] A. Kong, D. Zhang, and G. Lu, "A study of identical twins palmprint for personal verification," *Pattern Recognition*, vol. 39, no. 11, pp. 2149–2156, 2006. [35](#)
- [60] G. Lu, D. Zhang, and K. Wang, "Palmprint recognition using eigenpalms features," *Pattern Recognition Letters*, vol. 24, no. 9, pp. 1463–1467, 2003. [35](#)
- [61] X. Wu, D. Zhang, and K. Wang, "Fisherpalms based palmprint recognition," *Pattern Recognition Letters*, vol. 24, no. 15, pp. 2829–2838, 2003. [35](#)
- [62] G. M. Lu, K. Q. Wang, and D. Zhang, "Wavelet based independent component analysis for palmprint identification," *Proceedings of International Conference on Machine Learning and Cybernetics*, vol. 6, pp. 3547–3550, 2004. [35](#)
- [63] R. Chu, Z. Lei, Y. Han, and S. Z. Li, "Learning gabor magnitude features for palmprint recognition," *ACCV*, pp. 22–31, 2007. [35](#)
- [64] X. Y. Jing and D. Zhang, "A face and palmprint recognition approach based on discriminant dct feature extraction," *IEEE Transactions on Systems, Man and Cybernetics-Part B*, vol. 34, no. 6, pp. 2405–2415, 2004. [35](#)
- [65] M. Ekinici and M. Aykut, "Palmprint recognition by applying wavelet subband representation and kernel pca," *Lecture notes in Artificial Intelligence*, pp. 628–642, 2007. [35](#)
- [66] C. Han, H. L. Cheng, C. Lin, and K. Fan, "Personal authentication using palm-print features," *Pattern Recognition Letters*, vol. 36, no. 2, pp. 371–381, 2003. [35](#)
- [67] X. Wang, H. Gong, H. Zhang, B. Li, and Z. Zhuang, "Palmprint identification using boosting local binary pattern," *Intl. Conf. on Pattern Recognition*, pp. 503–506, 2006. [35](#)
- [68] K. Ito, T. Aoki, H. Nakajima, K. Kobayashi, and T. Higuchi, "A phase-based palmprint recognition algorithm and its experimental evaluation," *Proceeding of International Symposium on Intelligent Signal Processing and Communications*, pp. 2669–2672, 2006. [35](#)

- [69] X. Zhou, Y. Peng, and M. Yang, "Palmprint recognition using wavelet and support vector machines," *The ninth Pacific Proceeding of IEEE International Conference on Artificial Intelligence*, pp. 285–393, 2006. 36
- [70] http://www.lenovo.com/news/in/en/2006/12/lenovo_advanced_notebook.html. 37
- [71] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla, "Model-based hand tracking using a hierarchical bayesian filter," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1372–1384, 2006. 38, 54
- [72] J. Doublet, O. Lepetit, and M. Revenu, "Hand detection for contact less biometrics identification," *Cognitive Systems with Interactive Sensors, (2006)*, 2006. 38
- [73] V. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, pp. 677–695, 1997. 38
- [74] J. Doublet, O. Lepetit, and M. Revenu, "Contact less hand recognition using shape and texture features," *ICSP Proceedings 3, (2006)*, 2006. 39, 50, 54
- [75] M. Soriano, B. Martinkauppi, S. Huovinen, and M. Laaksonen, "Skin detection in video under changing illumination conditions," *International Conference On Pattern Recognition*, pp. 1839–1842, 2000. 40
- [76] V. Lepetit¹ and P. Fua², "Monocular model-based 3d tracking of rigid objects: A survey," *Foundations and Trends in Computer Graphics and Vision*, pp. 1–89, 2005. 41
- [77] W. A. Hoff, K. Nguyen, and T. Lyon, "Computer vision-based registration techniques for augmented reality," *Proceedings of Intelligent Robots and Control Systems XV, Intelligent Control Systems and Advanced Manufacturing*, pp. 538–548, 1996. 41, 43
- [78] J. Rekimoto, "Matrix: A realtime object identification and registration method for augmented reality," *Asia Pacific Computer Human Interaction*, 1998. 41
- [79] Y. Cho, W. Lee, and U. Neumann, "A multi-ring color fiducial system and intensity-invariant detection method for scalable fiducial-tracking augmented reality," *International Workshop on Augmented Reality*, 1998. 41
- [80] A. State, G. Hirota, D. Chen, W. Garrett, and M. Livingston, "Superior augmented reality registration by integrating landmark tracking and magnetic tracking," *Computer Graphics, SIGGRAPH Proceedings*, pp. 429–438, 1996. 41
- [81] C. Harris, "Tracking with rigid objects," *MIT Press*, 1992. 42
- [82] E. Marchand, P. Bouthemy, and F. Chaumette, "A 2d-3d model-based approach to real-time visual tracking," *Image and Vision Computing*, vol. 19, no. 13, pp. 941–955, 2001. 42
- [83] T. Drummond and R. Cipolla, "Real-time visual tracking of complex structures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 932–946, 2002. 42

- [84] A. Comport, E. Marchand, and F. Chaumette, “A real-time tracker for markerless augmented reality,” *International Symposium on Mixed and Augmented Reality (Tokyo, Japan)*, 2003. [42](#)
- [85] L. Vacchetti, V. Lepetit, and P. Fua, “Combining edge and texture information for real-time accurate 3d camera tracking,” *International Symposium on Mixed and Augmented Reality (Arlington, VA)*, 2004. [42](#)
- [86] D. G. Lowe, “Robust model-based motion tracking through the integration of search and estimation,” *International Journal of Computer Vision*, vol. 8, no. 2, pp. 113–122, 1992. [42](#)
- [87] D. Gennery, “Visual tracking of known three-dimensional objects,” *International Journal of Computer Vision*, vol. 7, no. 1, pp. 243–270, 1992. [42](#)
- [88] D. Koller, K. Daniilidis, and H. Nagel, “Model-based object tracking in monocular image sequences of road traffic scenes,” *International Journal of Computer Vision*, vol. 10, pp. 257–281, 1993. [42](#)
- [89] A. Kosaka and G. Nakazawa, “Vision-based motion tracking of rigid objects using prediction of uncertainties,” *International Conference on Robotics and Automation (Nagoya, Japan)*, pp. 2637–2644, 1995. [42](#)
- [90] A. Ruf, M. Tonko, R. Horaud, and H. Nagel, “Visual tracking by adaptive kinematic prediction,” *Proceedings of International Conference on Intelligent Robots and Systems*, pp. 893–898, 1997. [42](#)
- [91] B. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” *International Joint Conference on Artificial Intelligence*, pp. 674–679, 1981. [42](#)
- [92] M. J. Black and Y. Yacoob, “Recognizing facial expressions in image sequences using local parameterized models of image motion,” *International Journal of Computer Vision*, pp. 23–48, 1997. [42](#)
- [93] H. L. a nd P. Roivainen and R. Forchheimer, “3-d motion estimation in model based facial image coding,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 545–555, 1993. [42](#)
- [94] D. DeCarlo and D. Metaxas, “Optical flow constraints on deformable models with applications to face tracking,” *International Journal of Computer Vision*, vol. 38, pp. 99–127, 2000. [42](#)
- [95] M. Haag and H. Nagel, “Combination of edge element and optical flow estimates for 3-d model-based vehicle tracking intraffic image sequences,” *International Journal of Computer Vision*, vol. 35, pp. 295–319, 1999. [42](#)
- [96] G. Hager and P. Belhumeur, “Efficient region tracking with parametric models of geometry and illumination,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 10, pp. 1025–1039, 1998. [42](#)

- [97] F. Jurie and M. Dhome, “Hyperplane approximation for template matching,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 996–1000. 42, 45
- [98] F. Jurie and M. Dhome, “A simple and efficient template matching algorithm,” in *International Conference on Computer Vision*, 2001. 42
- [99] H. Moravec, “Towards automatic visual obstacle avoidance,” *International Joint Conference on Artificial Intelligence*, 1997. 43
- [100] W. Forstner, “A feature-based correspondence algorithm for image matching,” *International Archives of Photogrammetry and Remote Sensing*, vol. 26, no. 3, pp. 150–166, 1986. 43
- [101] C. Harris and M. Stephens, “A combined corner and edge detector,” *Fourth Alvey Vision Conference, Manchester*, 1988. 43
- [102] J. Shi and C. Tomasi, “Good features to track,” *Conference on Computer Vision and Pattern Recognition*, 1994. 43
- [103] S. K. Nayar, S. A. Nene, and H. Murase, “Real-time 100 object recognition system,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 12, pp. 1186–1198, 1996. 43
- [104] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” *Conference on Computer Vision and Pattern Recognition*, pp. 511–518, 2001. 43
- [105] F. Jurie, “Tracking objects with a recognition algorithm,” *Pattern Recognition Letters*, vol. 34, no. 19, pp. 331–340, 1998. 43
- [106] D. G. Lowe, “Fitting parameterized three-dimensional models to images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 441–450, 1991. 43
- [107] C. Schmid and R. Mohr, “Local grayvalue invariants for image retrieval,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 530–534, 1997. 43
- [108] T. Tuytelaars and L. VanGool, “Wide baseline stereo matching based on local, affinely invariant regions,” *British Machine Vision Conference*, pp. 412–422, 2000. 43
- [109] F. Mindru, T. Moons, and L. VanGool, “Recognizing color patterns irrespective of viewpoint and illumination,” *Conference on Computer Vision and Pattern Recognition*, pp. 368–373, 1999. 43
- [110] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 20, no. 2, pp. 91–110, 2004. 43
- [111] K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” *Conference on Computer Vision and Pattern Recognition*, pp. 257–263, 2003. 43
- [112] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: active contour models,” *International Journal of Computer Vision*, pp. 321–331. 43

- [113] A. Yuille, D. Cohen, and W. Hallinan, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, vol. 8, no. 2, pp. 99–111, 1992. [44](#)
- [114] D. Metaxas, "Physics-based deformable models," *Boston: Kluwer Academic*, 1997. [44](#)
- [115] S. Z. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near infrared images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, 2007. [45](#), [46](#)
- [116] A. Shashua and T. Raviv, "The quotient image: Class based rendering and recognition with varying illuminations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 129–139, 2001. [45](#), [46](#)
- [117] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001. [45](#)
- [118] R. Ramamoorthi, "Analytic pca construction for theoretical analysis of lighting variability in images of a lambertian object," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, pp. 1322–1333, 2002. [46](#)
- [119] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 218–233, 2003. [46](#)
- [120] H. Wang, S. Li, and Y. Wang, "Generalized quotient image," *IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 498–505, 2004. [46](#)
- [121] A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain, "Personal verification using palmprint and hand geometry biometric," *AVBPA*, pp. 668–678, 2003. [48](#), [49](#), [55](#)
- [122] G. Zheng, C.-J. Wang, and T. E. Boult, "Application of projective invariants in hand geometry biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 4, pp. 758–768, 2007. [48](#), [50](#), [51](#), [52](#)
- [123] C. Han, H. L. Cheng, C. L. Lin, and K. C. Fan, "Personal authentication using palmprint features," *Pattern Recognition*, vol. 36, pp. 371–381, 2003. [48](#), [49](#)
- [124] W. Shi, G. Rong, Z. Bain, and D. Zhang, "Automatic palmprint verification," *International Journal on Image and Graphics*, vol. 1, pp. 135–152, 2001. [49](#)
- [125] D. Zhang and W. Shu, "Two novel characteristics in palmprint verification: datum point invariance and line feature matching," *Pattern Recognition Letters*, vol. 32, no. 4, pp. 691–702, 1999. [49](#)
- [126] J. Chen, C. Zhang, and G. Rong, "Palmprint recognition using creases," *Proceedings of International Conference on Image Processing*, pp. 234–237, 2001. [49](#)
- [127] X. Lu, D. Zhang, and K. Wang, "Fisherpalms based palmprint recognition," *Pattern Recognition Letters*, vol. 24, pp. 2829–2838, 2003. [49](#)

- [128] G. Lu, D. Zhang, and K. Wang, “Palmprint recognition using eigen-palm like features,” *Pattern Recognition Letters*, vol. 24, pp. 1473–1477, 2003. [49](#)
- [129] Z. Sun, T. Tan, Y. Wang, and S. Z. Li, “Ordinal palmprint representation for personal identification,” *Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 279–284, 2005. [49](#)
- [130] A. Kumar and D. Zhang, “Personal authentication using multiple palmprint representation,” *Pattern Recognition*, vol. 38, no. 10, pp. 1695–1704, 2005. [49](#)
- [131] W. Li, D. Zhang, and Z. Xu, “Palmprint identification by fourier transform,” *Int. J. Pattern Recognition Artif. Intell.*, vol. 16, pp. 417–432, 2003. [49](#)
- [132] L. Zhang and D. Zhang, “Characterization of palmprints by wavelet signatures via directional context modeling,” *IEEE Transactions Sys. Man Cybern.*, pp. 1335–1347, 2004. [49](#)
- [133] A. Kumar and D. Zhang, “Integrating shape and texture for hand verification,” *Proceeding of the Intl. Conference on Image and Graphics*, pp. 222–225, 2004. [49](#)
- [134] D. Zhang and W. Shu, “Two novel characteristics in palmprint verification: Datum point invariance and line feature matching,” *Pattern Recognition*, vol. 32, no. 4, pp. 691–702, 1999. [49](#)
- [135] R. Hartley and A. Zisserman, “Multiple view geometry in computer vision,” 2000. [52](#), [53](#), [78](#)
- [136] L. Xiaowei, L. Yue, W. Yongtian, and Y. Dayuan, “Computing homography with ransac algorithm: a novel method of registration,” *Proceedings of the SPIE*, vol. 5637, pp. 109–112, 2005. [52](#), [54](#)
- [137] S. Garg, J. Sivaswamy, and S. Chandra, “Unsupervised curvature-based retinal vessel segmentation,” *Proc. of IEEE International Symposium on Bio-Medical Imaging(ISBI)*, pp. 344–347, 2007. [53](#), [67](#)
- [138] L. He and C. Xu, “Color photo denoising via hue, saturation and intensity diffusion,” pp. 159–169, 2008. [55](#)
- [139] C. Poon, D. Wong, and H. Shen, “A new method in locating and segmenting palmprint into region-of-interest,” *Proc. of 17th International Conference on Pattern Recognition(ICPR’04)*, pp. 1051–1054, 2004. [56](#)
- [140] O. Arandjelović and R. Cipolla, “A manifold approach to face recognition from low quality video across illumination and pose using implicit super-resolution,” *ICCV*, 2007. [63](#), [65](#)
- [141] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, “Advances and challenges in super-resolution,” *International Journal of Imaging Systems and Technology*, vol. 14, pp. 47–57, 2004. [64](#)
- [142] B. Zitová and J. Flusser, “Image registration methods: a survey,” *IVC*, vol. 21, pp. 977–1000, 2003. [65](#)

- [143] P. Gut, L. Chmielewski, P. Kukolowicz, and A. Dabrowski, “Edge-based robust image registration for incomplete and partly erroneous data,” *CAIP*, pp. 309–316, 2001. [66](#), [67](#)
- [144] G. Borgefors, “Distance transformations in digital images,” *Comput. Vision Graph. Image Process.*, vol. 34, no. 3, pp. 344–371, 1986. [66](#)
- [145] J. R. Movellan, “Tutorial on gabor filters,” [76](#)
- [146] <http://vision.ece.ucsb.edu/zuliani/Research/RANSAC/docs/RANSAC4Dummies.pdf>. [79](#)