Take up an idea, devote yourself to it, struggle on in patience, and the sun will rise for you.

–Swami Vivekananda

# Introduction

## 1.1 Preamble

In the near future, we are likely to encounter situations where human-computer interaction plays a vital role even in everyday menial tasks. Perhaps we are getting closer to a fully automated environment, where we frequently interact more with mechanical agents (such as ATMs, franchise voting machines, restricted access to buildings, laptops etc) than with our human counterparts. It is therefore the identity/security that has become an issue of overriding importance in 21st century. For these reasons, biometric based person identification or authentication has begun to grow rapidly and undoubtedly continue to do so, where fraud due to vulnerable security is costing public billions of rupees every year.

In the last several years, face recognition based person identification/verification has gained paramount importance predominantly due to the recent terrorist attacks. Although face has several advantages over other biometric traits such as fingerprint, hand geometry, iris, signatures etc, this problem of automatic face recognition is still largely unresolved although research has been conducted for several decades now. In this thesis, we have developed some novel algorithms based on subspace analysis for efficient and robust face recognition tasks. In addition, the proposed algorithms are tested using standard databases and their
performances are compared against well known existing algorithms as part of evaluation process. This helps to understand fundamental capabilities and limitations of algorithms, which is a key for developing new and efficient algorithms. In the following sections, we present overview of face recognition, its applications, technical challenges and current work. Later on, we present the motivations for the current work followed by an overview of contributions and finally we present the thesis outline.

1.2 Face Recognition

In this section, we present a brief outline of the face recognition research, its applications and technical challenges. We will also describe its contribution as a successful biometric trait.

1.2.1 What is Face Recognition?

Face recognition is a process of biometric identification by scanning a person’s face and matching it against a library of known or labeled faces [47]. In other words, given a still image or video segment, the task of face recognition is to identify one or more persons present in the scene. This task of face recognition is performed by humans in their daily activities routinely and effortlessly. The human brain perceives a person’s face quickly and uses that knowledge to recognize them under highly unpredictable and varied circumstances later. However, making a machine to perform this task with equal potential as that of human is still a herculean task though it has been studied for several decades now [50].

As one of the most successful application in the areas of image analysis and understanding, face recognition has received tremendous attention owing to the rapid growth of security based applications, especially during the last decade. This is highly evidenced by the emergence of dedicated conferences such as AFGR\(^1\), Sinobiometrics, and systematic empirical evaluation of face recognition techniques.

\(^1\)http://www.afgr2008.nl
including FERET[45], FRVT2, XM2VTS protocols3. In fact, research in face recognition is not only motivated by the technical challenges it poses, but also due to a wide range of applications it offers ranging from human-computer interaction to security and surveillance systems. Notwithstanding, there are at least three reasons for wide spread research on face recognition technology:

1. Availability of low-cost desktops, embedded computing systems, and feasible technologies to implement and realize practical face recognition systems.

2. Rapid growth of law enforcement and surveillance applications.

3. Perhaps, face is widely recognized as the feature which best distinguishes a person, often at-a-glance.

A face recognition system is a computer-driven application for automatically identifying a person from a digital image. It does this so by comparing extracted facial features in the live image with that of one from face database. This system is expected to identify faces present in images and as well videos. A generic face recognition system consists of 3 modules as depicted in Fig.1.1, where module-1 involves preprocessing steps such as face detection or face tracking, face alignment before face recognition. Module-2 is where the actual face recognition task is carried out and finally module-3 involves many post processing task such as facial expression analysis, face verification etc. We now explain briefly each steps of an automated face recognition system:

---

2http://www.frvt.org
3http://www.ee.surrey.ac.uk/Research/VSSP/xm2vtsdb
• **Face detection:** This is the first step in a face recognition system and its operational efficiency greatly impacts the success of face recognition task. Given a single image (or a video), the ideal face detector has to locate potential facial regions and segment them from non-face regions (or to track them in case of video) irrespective of their position, scale, and orientation. While a face detector provides a coarse estimate of the possible face regions, the face alignment component more accurately localizes and normalizes the faces. More information about this topic is available in [35].

• **Face recognition (or face identification):** This is a visual pattern recognition problem. It is a process of biometric identification by scanning a person’s face and matching it against a library of known or labeled faces. This involves a step called feature extraction, where the objective is to extract features that are useful for distinguishing faces of different persons. These features are expected to be invariant to variations such as age, expression, lighting etc. Comprehensive details about this topic is in [47].

• **Facial expression analysis:** Facial expression refers to the facial changes of a person’s internal emotions or intentions. Facial expression analysis refers to computer system that automatically analyze and recognize facial emotions from visual information. More information about this topic is available in [8].

• **Face verification (face authentication):** The objective of face verification is to validate a claimed identity based on the image of a face, i.e. either to accept or reject the claimed identity. This is one-to-one matching process unlike the face recognition task that involves one-to-many matching.

• **Watch List:** In Face Recognition Vendor Test 2002, which was conducted by the National Institute of Standards and Technology (NIST), another scenario was considered namely watch list. This is like answering to the query *Are you looking for me?* The test subject may or may not be present
in the database. The objective is to compute the similarity between the test sample and the library of labeled faces, and then to rank them in descending order of similarity score. If a similarity score is higher than a preset threshold value, an alarm would be raised to indicate the presence of test sample in the system’s database.

This thesis focuses on feature extraction models for efficient and robust face recognition tasks.

1.2.2 Face as a Biometric Trait

In the recent times, biometrics has become a field of extensive research due to its wide applications in authentication, security and so on. Biometrics can precisely be defined as the automatic recognition of a person by measuring features of biological traits. Biometrics enable identification of a person based on physiological and biological characteristics. Physiological biometrics are traits that are naturally grown (eg: face, fingerprint, body odour, DNA, etc), while behavior biometric traits are the ones which are learned or acquired (eg:face, signatures, gait, handwriting etc). When a system is using biometric verification system, the password to that will be feature of the human body itself. Some unique features that can act as passwords are iris, retina, face, voice, fingerprint and signatures.

Face recognition is widely accepted in many public sector scenario as most efficient biometric trait for the following key advantages:

1. One primary reason is it is non-intrusive in nature. While most biometrics require some degree of user interaction in order to acquire the biometric data, accurate face recognition can be performed by simply glancing the camera at the person from a distance.

2. Due to the non-intrusive nature, the public is likely to co-operate with the system to a great extent. It is well known that people fear for what they do not understand.
3. Face biometric is a unique one, because face is the only biometric belonging to both physiological and behavioral categories [67].

4. Implementation of face biometry is easy in any large organization. For instance, if a police department has to implement a iris or fingerprint biometric system, it would take years just to collect the data.

5. Perhaps the greatest advantage is that the biometric data required for identification is recognizable by humans. This would facilitate any manual check for identification or verification result, if system were to fail. Whereas any decision made by other biometric system such as iris or fingerprint, would require an expert to provide any reliable confirmation.

6. Unlike other biometric traits, face is the only biometric which can be used for general surveillance for its ability to collect the signatures from subjects without their active involvement.

1.2.3 Applications

Face recognition system are no longer limited to verification and surveillance tasks. Growing number of applications have started to use face recognition as the initial step towards interpreting human actions, intentions and behavior, as a central part of next-generation’s smart environment. Perhaps the key reason for face recognition to attract ample attention and to develop tremendously over the decades is its application in several government and commercial departments. In this section, we briefly explain some popular application of face recognition. Refer to [50, 62] for more review on face recognition applications.

**Face Identification:** Face recognition system identifies a person using their face images. This demands the presence of an authorized person instead of just validating an identification or PIN numbers, which in turn avoids the misuse of forgotten passwords and stolen cards.
**Access Control:** In scenario where less number of people are involved, such as office access or computer log-in, face recognition based system can achieve high accuracy without the need of much co-operation from the users. When this is combined with other biometric traits such as finger print or voice, much higher accuracy can be achieved.

**Security:** Today security is primacy, predominantly at airports, due to the recent terrorist attacks. Many airports around the globe have implemented face recognition technology. However, the desired level of low false reject and user satisfaction level are yet to be achieved due to the large number of faces that are to be examined at public places.

**Law enforcement:** Face recognition helps the investigators to identify suspects with incomplete information or even with rough sketch. Though the system performance is low due to the non-availability of the good quality face images of criminals, large number of police departments are being assisted by automatic face recognition systems.

**Multimedia:** Human faces are frequently encountered in news, sports, films, and other multimedia content. Indexing these multimedia content is important to generate video browsing, summarization etc. When combined with speech recognition, Natural Language Processing and other image understanding techniques, face processing becomes a powerful tool for ever-growing multimedia applications.

**Content based image retrieval:** Content based image retrieval, which resolves the difficulties encountered by text-based image retrieval, uses visual content such as colors, textures as features to index and retrieve the query images from the large databases. These common features have their own drawbacks, however, when they are combined with other image understanding techniques such as face recognition, the retrieval accuracy can be significantly increased.
Human-computer interaction: For an efficient and user friendly human-computer interactions, face emerges as the natural choice due to their capability in representing the exact mood or intentions about users. Indeed, this motivation could be one of the reasons for pioneering research on tracking and analyzing of human faces.

Apart from aforementioned applications, there are few other applications, which are potentially expected to use face recognition in near future. Those are antique photo verification and face image transmission[50].

1.2.4 Challenges

Although it is very easy for human beings to track, detect and recognize faces in complex background, building an automatic face recognition system remains very difficult because the face images can vary considerably in terms of facial expression, age, image quality and photometry, pose, occlusion and disguise [47]. In this section, we briefly describe the technical challenges, which are largely unresolved even with many state-of-the-art algorithms:
1. **Large variability due to intra-subject variations:** The appearance of a same face image may vary due to several factors such as variations in expression, lighting, pose, time and also due to occlusion. Fig. 1.2 shows an example of intra-subject variations due to these factors. It clearly shows the complexity involved in carrying out face-based person-identification task. Hence, the variations between the images of the same face due to illumination and viewing directions are almost larger than the image variations due to change in face identity.

2. **Small Sample Size problem:** The number of samples per person available for training the underlying manifold is usually much lesser than the dimensionality of the image space. Under such conditions, the system may not generalize well to recognize unseen samples of the face.

3. **Single sample per person for training:** Ability to work under single training sample per person requires tremendous care and attention in order to achieve desired level of accuracy. Many successful face recognition algorithms developed today heavily depends on the representative training samples and even some algorithms completely fail under this condition [52].

4. **Unconstrained conditions:** Many algorithms reported in literature perform well only under highly controlled scenarios. Their performance significantly deteriorates when confronted with unconstrained situations [67]. Unconstrained conditions include uncontrolled background, video sequence, variations in face due to aging, make-up and cosmetics, noise etc.

5. **Subject’s non-cooperation:** Finally, also due to the subject’s non-cooperation such as the refusal to look at the camera and moving subject also greatly hinder the performance of the face recognition system.

Even though substantial progress in face recognition has been realized, a solitary efficient algorithm which can cope up with aforementioned difficulties is yet to be developed.
1.3 Literature Survey

In recent years, face recognition has received substantial attention from both research communities and the market, but still remained very challenging in real applications. A lot of face recognition algorithms, along with their modifications, have been developed during the past decades. The contemporary face recognition algorithms can mainly be classified into two categories:

1. **Model-based schemes**: which uses shape and other texture of the face, along with 3D depth information.

2. **Appearance-based schemes**: which uses the holistic texture features.

1.3.1 Model-Based Schemes

It is important to have a model of face that can capture variation in face appearance in order to interpret images of faces. The changes that can happen can essentially be in two parts: changes in shape and changes in texture pattern across the face. Both shape and texture can also vary because differences between individual and also due to changes in expression, lighting, viewpoint variations. There exists a strong concept known as model based approaches (statistical models of appearance), that will generate compact models of shape and texture variation to interpret images of faces.

In these approaches, a statistical learning model is adopted that can capture the ways in which the shape and texture of faces vary across a range of faces. The approach relies on a large and representative training set of facial images, each of which is annotated with a set of feature points (landmark points). The positions of these landmark points are used to define the shape of the face and its variation across the training faces. The pattern of intensities are then analyzed to learn the ways in which texture can vary. These two models together is capable of synthesizing any of the training images and generalize them to interpret the new faces. The prior knowledge of human face is utilized to design the model.
One of the earliest face recognition algorithms based on automatic feature extraction was proposed by Kanade [33]. This system localizes the corners of eyes, nostrils, etc. in frontal view images, and compares the features against parameters of known faces (Using Euclidean metric). A recent feature-based system, based on elastic bunch graph matching, was developed by Wiskott et al. [34] as an extension to their original graph matching system [36]. Cootes et al developed a 2D morphable face model, through which the face variations are learned [56]. Many extended algorithms based on this concept have been proposed. In our research, we concentrate on alternate approaches known as subspace methods (also known as appearance based schemes), which is more simple (in the sense that it does not involve the manual landmarking of facial features as in case of model based schemes) yet effective as it can withstand large variations in images.

1.3.2 Appearance-Based Schemes

This section provides a brief overview of subspace based approaches that are successfully developed for face recognition. The objective of subspace based approach is to project the data of faces onto a dimensionally reduced space where the actual recognition will be carried out. Turk and Pentland in 1991, first explored the Principal Component Analysis (PCA) [38] for face recognition and used the PCA projected components as the features. The PCA is an unsupervised learning technique and hence does not include the label information of the data.

In order to utilize the class label information of the data, Linear Discriminant Analysis (LDA) was proposed [46, 28]. This technique computes the basis vectors from the underlying data that optimally discriminates among classes. This is unlike the PCA method, which searches for basis vectors that best describes the data\(^4\). The objective of LDA is to maximize the between-class measure while minimizing the within-class measure. However, due to large dimensions, implementation of the LDA method becomes an intractable task, because of singularity of within-class matrix [5]. To resolve this, the original \(n\) dimension of the data

\(^4\)More information about relative performance between PCA and LDA is available in [5]
is projected onto $l$ dimensional space using PCA, where $l \ll n$. This PCA+LDA representation is known as Fishers LDA (FLD) or Fisherface method[43].

The concept of Independent Component Analysis (ICA) was also explored for face recognition [37]. The ICA is a generalization of the PCA algorithm which extracts the information contained in the higher-order relationships among pixels also. It differs from PCA in the following three aspects: i) The basis vectors obtained through ICA need not be orthogonal in nature, ii) The ICA can also handle data which are non-gaussian in nature and iii) the ICA minimizes the higher-order dependencies, unlike the PCA which minimizes the second order(moments) dependencies of the data. The ICA was investigated with two fundamentally different architectures: Architecture-1 (ICA-I) and Architecture-2 (ICA-II). Where ICA-I reflect more local properties of the faces while the ICA-II reflect more global properties and thus closely captures the face information.

As an alternative to the Principal Component Analysis, the Locality Preserving Projections (LPP) also known as Laplacian faces, was proposed which optimally preserves the neighborhood structure of the data set [22]. The LPP shares many of the data representation properties of nonlinear techniques such as Laplacian Eigenmaps or Locally Linear Embedding.

These algorithms are state-of-the-art subspace methods proposed for face recognition (See Ref.[18] for a more detailed description about PCA, LDA, ICA etc.). Many variants of these algorithms are devised to overcome specific anomaly such as storage burden, computational complexity, Single Sample Per Person (SSPP) problem etc. Especially under latter circumstance (i.e.SSPP), performance of many algorithms deteriorate due to inadequacy of representative training samples. Fortunately, several ways to get around this difficulty have been proposed [61, 12, 65, 4, 63]. Couple of attempts were also made to make the Fisherface method compatible under SSPP problem [64, 11, 26]. In addition, many extended versions of these algorithms have been proposed based on the standard eigenface method during the last several years, such as probabilistic eigenface method [9], Support Vector Machines (SVM) method [44], feature-line method[49], evo-
Correlation techniques for improving the classification performance and also for reducing the computational complexity of the subspace based face recognition have also been presented [57].

To combine the merits of PCA, LDA and Bayesian subspace approaches, a single framework model was presented in [58]. Here a 3D parameter space was constructed using the three subspace dimension as axes. Much improved accuracy was achieved through this unified subspace approach and which does not suffer from the individual limitations. To enhance the performance of PCA and LDA methods, instead of extracting single set of features, more than one set of features were extracted by using Gaussian Mixture Models [29, 30, 18]. These models, though computationally expensive, obtained good performance under large variations of pose and/or illumination. Further, in order to utilize the contribution made by the local parts of the whole image, the PCA and FLD methods are applied component-wise on the whole image [13, 39]. This has resulted in improved accuracy under varied facial expression and lighting configurations and justifies increased computational efforts.

Several attempts have been made to reduce the computational burden of the original PCA and FLD methods [14, 15]. In [14], the original PCA and LDA methods were implemented in the DCT domain and it has been theoretically proved that the results obtained in the DCT domain are exactly the same as the ones obtained in the time/spatial domain. Similarly, the PCA and FLD algorithms have been implemented in the wavelet domain [15]. The merit of transforming the original image from the time domain to frequency domain is that, in frequency domain we will be working on the reduced coefficient set which are more prominent for recognition task and it also helps in reducing the computational burden.

The standard PCA and FLD methods, though popular and widely admired, are computationally expensive, because, the 2D image has to be aligned into 1D vectors prior to feature extraction. This leads to large covariance matrix and consequently the eigen value decomposition task becomes tedious and time consuming. To resolve this, 2DPCA [27] and 2DFLD [33] were proposed which does
not involve the matrix to vector conversion\(^5\). Because of this, several appreciable merits were achieved: i) Computation of scatter matrix (or scatter matrices in case of Fisherface method) became simple ii) Consequently, computation time also reduced and iii) significant improvement in accuracy was achieved as compared with their 1D counterparts. However, the main drawback is that these methods require huge feature matrix for representation and recognition tasks. Later, even this drawback of 2DPCA and 2DFLD was resolved by 2D^2PCA \([20, 32, 70]\) and 2D^2LDA \([40]\) methods respectively. These methods are unlike their 2D counterparts \([27, 33]\) which used either the row or the column projections at a time. The recognition accuracy of these methods was either the same or even higher than the corresponding 2D versions, but their computation time and storage requirement was significantly reduced as compared with the 2DPCA and 2DFLD methods.

More recently methods namely DiaPCA \([66]\) and DiaFLD \([41]\) have been proposed. These methods seek optimal projection vectors from diagonal face image so that the correlation between variations across rows and columns of images can be preserved \([66]\). Substantial improvement in accuracy was achieved by DiaPCA and DiaFLD methods than the 2DPCA and 2DFLD methods respectively. Subsequently, these DiaPCA and DiaFLD methods were combined with 2DPCA and 2DFLD methods respectively, in order to achieve efficiency both in terms of accuracy and storage requirements.

Very recently, a significant improvement to subspace method has been proposed by taking a tensor approach to image data \([53, 54]\). Two crucial aspects of this has been rapid convergence (between 10 and 30 iterations) and as is keenly sought after in optimization techniques, the ability to get over local minima. These results have taken the standard supervised learning approaches of Support Vector Machines, minimax probability machine, FLD and extended in to tensored structure (multi linear form) of this with improved results. In terms of real time application, while the rapid convergence shown in these papers \([53, 55]\) is

\(^5\)Refs. \([23, 33]\) are two similar methods found in the literature referred respectively as 2DFLD and 2DLDA. In all our discussions, whenever we refer to Ref.\([33]\), it also implies Ref.\([23]\) and vice-versa.
encouraging, what is not clear is the computational complexity involved. This is because the authors have been using optimization methods based on idea of alternate projections, which is known to be computationally burdensome.

From the survey of above literature, it is clear that the subspace based algorithms have a consistent progress since last two decades and even now continue to yield newer and better results. In particular during the last five years, there has been expansion of investigation and implementation needs for real time video/streaming problems. This and other concerns are driving the research community to even focused activity now.

1.4 Motivation

Face recognition has innumerable of applications, including security, person verification, Internet communication, and computer entertainment. Even though research in automatic face recognition has been conducted since several decades, there exists many open problems which deserve further study. Systems have been developed for face detection and tracking, but reliable face recognition still offers a great challenge to computer vision and pattern recognition researchers. Thus, research in face recognition is not only motivated by the potential application it offers but also due to the inherent technical challenges involved.

In addition, from the survey of literature it is quite evident that subspace based algorithms for face recognition are widely admired and adapted in current face recognition research. The subspace methods, which are unlike the explicit facial features based methods, are being widely used now-a-days due to its insensitiveness to noise. The prime motivation for this thesis is to propose new learning subspace based methods to overcome the limitations of the existing algorithms.

In signal or image processing applications, there is a strong urge for the development of algorithms which can withstand any conceivable noise that can arise within the practical framework of the signal. However, many contemporary successful algorithms cannot successfully deal with noise conditions and thus their
performance drops significantly. In real time applications the issue of processing signals corrupted by noise of various nature dominate the utility of such algorithms. This we feel is more important while addressing real time applications. Consequently it becomes crucial to check the robust behavior of the algorithms under real time pattern recognition and computer vision tasks.

Thus our motivation for this work are threefold:

1. Potential applications and challenges involved in face recognition.
2. The concept of subspace analysis and its impact on face recognition research.
3. To study the robust nature of various subspace based algorithms.

1.5 Highlights of This Thesis

To highlight, following are the prime contributions of this work, which will be detailed in subsequent chapters:

1. We have proposed some novel subspace based algorithms for efficient and robust face recognition in three different domains (spatial, frequency and combined time-frequency).

2. To improve the recognition accuracy, we have justified the need for a mathematical process (Gram-Schmidt decomposition process) as a preprocessing step for FLD based algorithms.

3. Performance evaluation of about twenty different subspace algorithms under several real time test conditions.

4. Conducted extensive experiments to study the robustness of existing subspace algorithms under five different noise conditions. In addition, we have suggested robust algorithms and also compared them under noise conditions.

5. Provided analysis of experiments and substantiated through appropriate reasons based on the performance.
1.6 Organization of This Thesis

The subsequent chapters in the thesis have been divided into five main chapters: a chapter for describing tools and techniques used in our work, three chapters each for representing the contributions of the work in spatial domain, frequency domain and combined spatial-frequency domain respectively. Finally a chapter for presenting the consolidated results obtained as part of performance evaluation process.

In Chapter 2, we will present some tools and techniques that will be used in this work. The techniques discussed are fundamentals of subspace techniques and related tools like PCA, LDA etc., transform domain techniques such as DCT, Block DCT and Wavelets and few orthogonalization techniques like SVD, QR, Gram-Schmidt decomposition etc. We also give the notations and abbreviations used throughout the thesis.

In Chapter 3, we develop some spatial domain subspace algorithms based on FLD analysis for efficient and robust face recognition task. We have also emphasized the need for a preprocessing step for FLD based algorithms. In addition, we have proposed a method called LPP mixture model based on Gaussian Mixture Models. Finally, to study the robustness of subspace algorithms, we have modeled five different noise conditions using various continuous distributions and used them in their discretized version.

In Chapter 4, the potentiality of PCA mixture model [29] and LPP mixture models [43] are further explored by implementing them in frequency domains such as Discrete Cosine Transform (DCT). We conduct several experiments (by varying the percentage of DCT coefficients) under both clean and noise conditions and have presented the detailed analysis of the result obtained.

In Chapter 5, We indicate the subjective nature of wavelets. We also propose two methods based on mixture models using Wavelet coefficients. The advantage of both spatial and frequency components are captured by these mixture models through the application of Wavelets.
In Chapter 6, a comprehensive performance evaluation of around twenty five different subspace algorithms are carried out under many important real time test conditions. Some crucial theoretical aspects of algorithms and analysis of experiments are presented in detail.

In Chapter 7, we summarize the major contributions of this thesis. After presenting concluding remarks, we indicate some possible future avenues that one can pursue based on this research.