Literature Assessment for Pose-Invariant Face Recognition

S. Jebha, P. KANNAN
PG Scholar Department of ECE, PET Engineering College, India
Professor Department of ECE, PET Engineering College, India

Abstract:
Face recognition becomes problematic when the variation is considered across different poses. To overcome this problem, invariant pose is considered. For this invariant pose, Optimization algorithm is used. Using this algorithm, frontal image is reconstructed virtually from the profile of non-frontal image. This literature survey gives in detail about the different kinds of methods which are related with variant poses.

Keywords: Invariant poses, Optimization algorithm, Fisherfaces, Appearance model, Morphable model, Light fields, Probabilistic stack-flow.

I. INTRODUCTION

Face recognition is used in many applications such as access control, identification systems, surveillance, biometrics and pervasive computing. In face recognition, pose variations is considered as one of the important problem, which would not produce better accuracy and performance. Face recognition across poses is carried out in 2D and 3D models. Many algorithms are implemented to compute 2D and 3D models to obtain better results. The performance degrades when there are variations in the face with pose, illumination and expression [2]. Face recognition algorithms such as Eigenfaces and Fisherfaces[1] are used under pose variations which produce low recognition rates only. This algorithm cannot be applied to 3D faces. Face recognition differ for both 2D and 3D techniques. The variations across pose and illuminations can be handled by the algorithm which helps to simulate the process of image formation in 3D space, using computer graphics, and it estimates 3D shape. This is achieved by using morphable model of 3D faces [4]. The stereo matching [22] technique in 2D image is to judge the similarity of two images that are viewed from different poses. This stereo matching helps for the purpose of arbitrary, physically valid and continuous correspondences. Gabor–Fisher Classifier (GFC) method [18] is robust to illumination and variability in facial expression. This method is applied to feature which can be extracted from the gabor filter through Enhanced Fisher linear discriminant Model (EFM). The patch transform [12] represents an image into overlapping patches that are sampled on a regular grid. This transform helps to influence images in the patch domain, which then begins the inverse patch transform to synthesize modified images. In tied factor analysis [19], the linear transformation depends on the pose, but the loadings are constant (tied) for a given image. So Expectation-Maximization (EM) algorithm is used for the estimation of linear transformations and noise parameters from training data.

II. LITERATURE REVIEW

A. Eigenfaces and Fisherfaces:
On seeing that correlation methods are computationally expensive and require enormous amount of storage, it is normal to go behind dimensionality reduction schemes. A technique that is commonly used for dimensionality reduction in computer vision mostly in face recognition is principal components analysis (PCA) [16]. This technique selects a linear projection and maximizes the scatter of all projected samples by dimensionality reducing it. If the image is linearly projected with eigenvectors of same dimension as in the original images, then it is known as Eigenfaces. Here, the classification is performed by means of a nearest neighbour classifier in the reduced feature space and then the Eigenface method is equivalent to the correlation method. A drawback of this method is that there is unwanted information in the scatter which is maximized not only to the between-class scatter that is helpful for classification, but also for the within-class scatter for classification purposes. Thus if PCA is applied to face images under varying illumination, the projection matrix will contain Eigenface which maintains the projectefeature space. The points that are present in the projected space will not be clustered and not as good as the classes that are grimed together.

The Eigenfacemethod has the advantage that, the variation within classes lies in a linear subspace of the image space. Consequently, the classes are convex and linearly separable. Dimensionality reduction can be done using linear projection and can conserve linear separability. But dimensionality reduction is still a problem in recognizing a face as it is insensitivity to lighting conditions. Here the learning set is labelled, it
so that the information can be used to reduce the dimension of the feature in an image. So, class specific linear methods can be used for dimensionality reduction and simple classifiers to reduce the feature space, as a result better recognition rates can be obtained than the Eigenface method. Fisher’s Linear Discriminant (FLD) [1] is used as a class specific method, it can shape the scatter in order to make it additional consistent for classification. The problem in face recognition is that it is difficult to maintain the within-class scatter matrix at all times singular. The scatter matrix is given below,

$$S_f = \sum_{k=1}^{N}(x_k - \mu)(x_k - \mu)^T$$  \hspace{1cm} (1)$$

Where \(X_k\) is the ‘N’ number of sample images and \(\mu\) is the mean images of samples. It is feasible to decide the within-class scatter matrix of the projected samples to compose it exactly to zero. In order to overcome the difficulty of a singular scatter matrix Fisherface method is used. It overcomes the problem by projecting the image set to a lower dimensional space and obtain the ensuing within-class scatter matrix as nonsingular. This is attained by using PCA to reduce the dimension of the feature space and then Fisherface method is applied to reduce the dimension into nonsingular matrix.

Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor have described a method for matching the deformable models for shape and appearance of original images. The advance of Edwards et al. [5], the statistical appearance models was generated from the combination of shape model and its variation with texture model. The term “texture” means the intensities in the pattern or the color transversely in the patch of an image. This model is assembled from the training set of annotated images in which analogous points are marked. For constructing a face model, the points are marked on the face image which defines the main features. Then Prerectes analysis is applied for the alignment of the sets of points in face which is represented as the vector \(x\), thus a statistical shape model is built. And then warp of the training image is matched with the image so that a “shape-free patch” is obtained. The appearance model has parameter \(c\) which controls the shape and texture in the model frame from the given equation below,

$$x = \bar{x} + Q_s c$$
$$g = \bar{g} + Q_g c$$  \hspace{1cm} (2)$$

Where, \(\bar{x}\) is the mean shape, \(\bar{g}\) is the mean texture in a mean shaped patch, \(g\) is the texture vector and \(Q_s\), \(Q_g\) are the matrices describing the modes of variation derived from the training set. To make a texture model Eigenface method is used by using PCA to reduce the dimension of the feature space and simple classifiers to reduce the feature space, as a result better recognition rates can be obtained than the Eigenface method. Fisher’s Linear Discriminant (FLD) [1] is used as a class specific method, it can shape the scatter in order to make it additional consistent for classification. The problem in face recognition is that it is difficult to maintain the within-class scatter matrix at all times singular. The scatter matrix is given below,

$$S_f = \sum_{k=1}^{N}(x_k - \mu)(x_k - \mu)^T$$  \hspace{1cm} (1)$$

From the above Fig 1, Fisherfacemethod has slow error rates than the Eigenface method and less computational time. The disadvantage of Fisherface method is that it produces very low recognition rates.

B. Appearance model:

Analysis applied to the image frame and is generated by applying scaling and offset to the intensities. Then correlations are made between shape and texture and a combined appearance model is generated. The reconstruction can be obtained by producing the texture in a mean shaped patch, then warping it consequently to facilitate the model points lay on the image points. The parameters of appearance model and shape transformation gives the shape of the image patch which are defined by the position of the model points in the image frame so as to represent the model. During matching, the pixels of the image are sampled and projected into the texture model frame. The elements present in a row of the matrix give the weights to each texture sample point for the model parameter. This method makes use of correlation between errors in model parameters and the remaining texture errors. The disadvantage of this method is that it cannot be used for tree like structures with varying numbers of branches, but it can be used for assorted shapes which exhibit shape variation but not a change in topology. AAM search is slightly slower than Active Shape Model.

C. Morphable model:

Morphable model of 3D faces captures the class-specific properties of faces. These properties are learned automatically from a data set of 3D scans. The
morphable model represents shapes and textures of faces as vectors in a high-dimensional face space, and involves a probability density function of natural faces within face space. The algorithm used here estimates all 3D scene parameters automatically, including head position and orientation, focal length of the camera, and illumination direction. This is achieved by a new initialization procedure that also increases robustness and reliability of the system considerably. The new initialization uses the image coordinates of between six and eight feature points. Figure 2

In morphable face model (Fig. 2), the recognition is based on fitting the intrinsic shape and texture of faces. This model is independent of imaging conditions. In the fitting algorithm, all the gallery images are analyzed for the purpose of identification. In the probe image, also fitting algorithm is computed to obtain the model coefficients. They are then compared with all the model coefficients obtained from gallery data to locate the nearest neighbour. The morphable face model [4] is based on a vector space representation of faces that is constructed such that any bowed combination of shape and texture vectors. Continuous changes in the model parameters generate a smooth transition such that each point of the initial surface moves toward a point on the final surface. Dense point-to-point correspondence is crucial for defining shape and texture vectors. 3D morphable model is prevailing and adaptable representation for human faces. But it requires many manually selected landmarks for initialization.

D. Eigen light-field (ELF) method:

The important attribute of appearance-based algorithms is based on the recognition decision which is obtained by considering the features as the pixel intensity values in an image. The pixel intensities which are considered as features in appearance-based algorithms match up straight to the radiance of light emitted from the object along certain rays in space. The light-field specifies the radiance of light along all rays in the scene. Hence, the light field of an object is the set of all possible features that could be used by an appearance-based algorithm. Any number of images can be used, from one upwards, in both the gallery and the probe sets. The light-field correctly renders images across pose. This algorithm can use any number of gallery images captured at arbitrary pose and any number of probe images also captured with arbitrary poses.

Fig 3. The average recognition rate is figured across gallery and probe poses using Eigen Light-Fields (ELF), FaceIt, and Eigenfaces.

This algorithm operates by estimating the light-field of the subject’s head. First, generic training data is used to compute an eigen-space of head light-fields, similar to the construction of Eigen-faces. The projection into the eigen-space is performed by setting up a least-squares problem and solving for the projection coefficients. This linear algorithm can be applied to any number of images, captured from any poses. Finally, matching is performed by comparing the probe and gallery light-fields using a nearest neighbour algorithm. Capturing the complete light-field of an object is a difficult task, primarily because it requires a huge number of images. From Fig 3, it is inferred that the average recognition rate by ELF method produces better recognition rate[6] than other method such as Eigenfaces and FaceIt. The Eigen Light-Field Estimation Algorithm can be used to estimate a light-field from a collection of images. Once the light-field has been estimated, it can then, theoretically at least, be used to render new images of the same object under different poses. To use it for face recognition across pose, it needs to perform vectorization, classification, selecting training and testing sets. Vectorization converts measurements of light into features for a pattern recognition algorithm. The Eigen Light-Fields Algorithm can be extended to recognize faces across pose and illumination simultaneously by generalizing eigen light-fields to Fisher light-fields.

E. Probabilistic stack-flow:

Stack flow ascertainment the viewpoint induced spatial deformities undergone by a face at the patch level. It helps to learn the correspondence between patches coming from two different viewpoints of a face. In this stack flow, a warp is consistent across the entire stack and aligns a patch in the stack of profile faces to a patch in the stack of frontal faces. It achieves resistance to noise by using the stack of images as a constraint and avoids the averaging process. By
avoiding the averaging, it is better to handle at texture variation. The discriminative power of each warped patch as a function of viewpoint, which is known as Probabilistic stack-flow [8]. The probability distribution for patch SSD is given as

\[ p(s_r | \phi_p, \omega), \omega \in \{c, l\} \quad (3) \]

where \( s_r \) is the SSD score for the \( r \)-th patch, \( \phi_p \) is the probe viewpoint and the variable \( \omega \) refers to classes when gallery and probe images belong to the same subject of client \( C \) or different subjects of impostor \( I \). Here, the sum of squared differences (SSD) is used as a similarity value between a gallery patch and a warped probe patch. The limitation with this method is that it is inability to handle situations where patches become occluded at extreme viewpoints.

F. Optimization algorithm:

Even though there are many methods for face recognition across poses, they all have low recognition rates and attain less efficiency. In order to overcome these problems, optimization algorithm [21] is used. This approach consists of two methods. The first one is the Markov Random Fields (MRF) and second one is the Belief Propagation (BP) algorithm. This approach can produce high recognition rates and increase in efficiency.

![Image](image.png)

Fig 4. Recognition rates of different multi-pie database at various viewpoints.

From the above Fig 4, it is inferred that for various methods the obtained recognition rates fluctuates. The method MRF and BP gives better recognition rate than other methods.

a) Markov Random Fields:

Markov random field models provide a robust and integrated framework for early vision problems such as stereo, optical flow and image restoration. MRF helps to find the globally optimal set of local warps that can be used to predict the image patches at the frontal view. In order to reduce the effect of illumination changes, the local patches are normalized by subtracting the means and dividing by the standard deviations before estimating the sum of squared difference in the overlapping region. The optimal set of warps is found by minimizing the energy function. After minimizing, the energy function is given as,

\[ E(|p_1|_4) = \sum_{i \in V} E_i(p_i) + \lambda \sum_{(i,j) \in E} E_{ij}(p_i, p_j) \quad (4) \]

Where \( p_i \) and \( p_j \) are the warps taken from two different nodes.

b) Belief Propagation:

Priority-BP can, thus, be viewed as a generic way for efficiently applying belief-propagation to MRFs with very large discrete state spaces, thus dealing, for the first time, with what was considered as one of the main limitations of BP up to now. At this point, it should be noted that priority-BP, as any other version of BP, is only an approximate global optimizer. Priority-belief propagation (BP) algorithm, which carries 2 major improvements over standard BP [10]: dynamic label pruning and priority-based message scheduling. Priority-BP is a generic algorithm, applicable to any MRF energy function. BP is an iterative algorithm that tries to find a MAP estimate by iteratively solving a finite set of equations until a fixed point is obtained. BP continuously propagates local messages between the nodes of an MRF graph. At every iteration, each node sends messages to all of its neighbouring nodes, while it also accepts messages from these nodes. During the forward pass vigorously reduces the number of possible labels for each node by discarding labels that are unlikely to be assigned to that node. After all MRF nodes have pruned their labels at least once then it can be precomputed by reducing the matrices of pair wise potentials and thus the speed of BP algorithm is increased. The use of message scheduling to determine the transmitting order for a node based on the confidence of that node about its labels. The node most confident about its label should be the first one to transmit outgoing messages to its neighbours.

III. CONCLUSION

In this paper, literature assessment for pose-invariant face recognition is discussed in concise. It is seen that
the optimization algorithm provides fine recognition rates than any other recognition methods.

ACKNOWLEDGEMENT

Apart from the efforts of me, the success of any work depends largely on the encouragement and guidelines of many others. I take this occasion to express my gratefulness to the people who have been influential in the successful completion of this work. I would like to extend my sincere thanks to all of them. I owe a sincere prayer to the LORD ALMIGHTY for his kind blessings and giving me full support to do this work, without which this would have not been possible. I wish to take this opportunity to express my gratitude to all, who helped me directly or indirectly to complete this paper.

REFERENCES


Fig 2. Morphable face model performed by the comparison of model coefficients.