Multi-Task Pose-Invariant Face Recognition

Digital Image Processing MATLAB Project
Aman Shenoy – 2016A8PS0393P

***Abstract*— Face images captured in unconstrained environments usually contain pose variation, which significantly degrades the performance of algorithms designed to recognize frontal faces. This project implements a novel face identification framework capable of handling pose variations within ±45° of yaw. The proposed framework first transforms the original pose-invariant face recognition problem into a partial frontal face recognition problem. A robust patch-based face representation scheme is then developed to represent the synthesized partial frontal faces. For each patch, a transformation dictionary is learnt under the proposed multitask learning scheme. The transformation dictionary transforms the features of different poses into a discriminative subspace.**

Finally, face matching is performed using a PCA-based Face Recognition system called 'Eigenface' where the synthesised frontal face is matched with the correct frontal face from within a frontal face dataset. Experimentation was done using captured image of a subject at multiple angles and the frontalization algorithm was tested on multiple images in the LFW (Labelled Faces in the Wild) dataset. Extensive and systematic experimentation done by the reference paper on FERET, CMU-PIE, and Multi-PIE databases shows that the proposed method consistently outperforms single-task-based baselines as well as state-of-the-art methods for the pose problem.

***Index Terms*—Pose-invariant face recognition, partial face** recognition, multi-task learning.

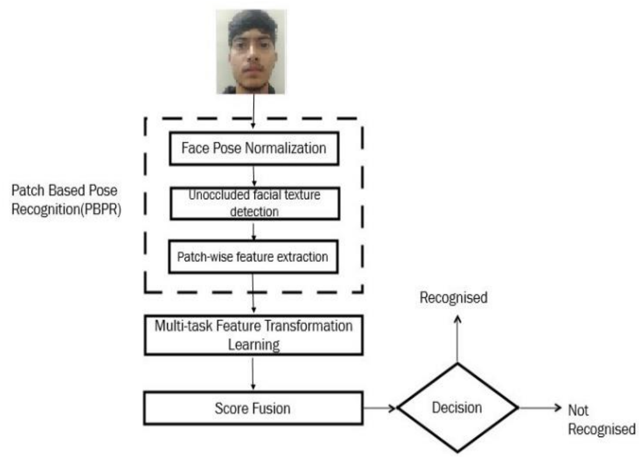
I. INTRODUCTION

Motivation - Face recognition has been one of the most active research Topics in computer vision for a long time. It is a challenging problem which has received much attention during the recent years due to its potential multimedia applications in different fields such as 3D videoconference, security applications or video indexing. Large variations of pose exhibited by the subjects in the stated applications has been a major challenge in the field of Computer Vision. Pose Invariant Face Recognition is indispensable in realizing the full potential of Facial Recognition

The presented approach mainly handles the identification problem of matching an arbitrary pose probe face with frontal gallery faces, which is the most common setting for both the research and application of pose-invariant face recognition (PIFR).

Pose problem usually combined with other factors, such as variations in illumination and expression, to affect the appearance of face images results into making an extent of appearance change caused by pose variation greater than that caused by differences in identity. Thus, the performance of frontal face recognition algorithms degrades dramatically when the images to be matched feature different poses. So, directly matching faces in different poses becomes difficult.

Proposed Method –



Basic Block Diagram of proposed method implemented in the project

Existing face representation tend to extract fixed length features from images with the assumption that all facial components are visible in the image. The PBPR scheme is flexible representation of the face where the length of the face is related to the pose.

The frontal face is generated using this leaving us with a transformation matrix for that transformation. Multiple images of a subject at various angles are then collected and a dictionary of transformation matrices is then formed in which each transformation matrix is now in a common discriminative subspace and hence can learn from each other by Multi-task Feature Transformation Learning. This helps in accounting for and modeling the correlation between the mutliple poses of the same person hence optimizing the transformation matrices for better frontaliation results.

MtFTL even though extremely powerful is not as practical as collecting images of the same person in multiple poses, to train the learning algorithm, is very difficult and impractical considering unconstrained environments. This project will mainly focus on frontalization and matching and will briefly touch upon Multi-Task Feature Transformation Learning due to lack of public availaibility of FERET and CMU-PIE datasets.

II. Implementation

Before discussing the implementation, the problem statement and objectives are made clear for the project

Problem Statement - Design and development of a pose invariant algorithm for face recognition with the help of multi task feature transformation learning.

Objectives –

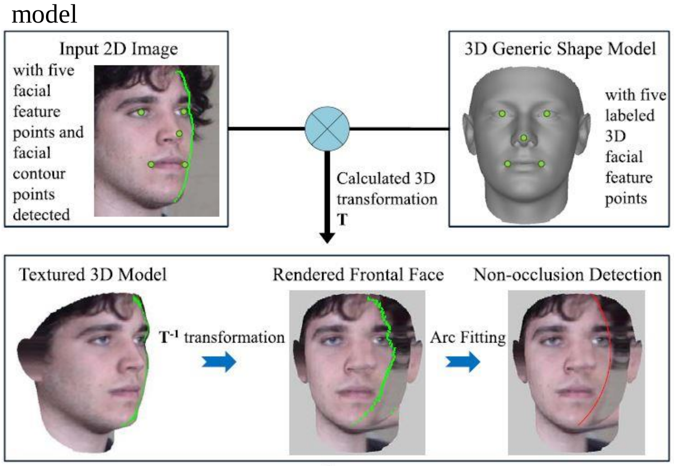
1.) Implement a working model for frontalization of face images with acceptable results up to ±45° of yaw
2.) Implement MtFTL on a set of images of a subject captured in different poses and obtain optimized transformation matrix dictionary
3.) Implement a face matching algorithm for matching synthesized frontal face with the correct frontal face with greater than 90% accuracy

The first step of generating a synthesized frontal face would be Patch Based Pose Recognition and Frontalization which consists of the Face Pose Normalization, Un-Occluded facial texture detection and Patch Based Face Representation. The theoretical aspects are described first followed by a discussion on how it has been implemented on MATLAB.

A. Face Pose Normalization

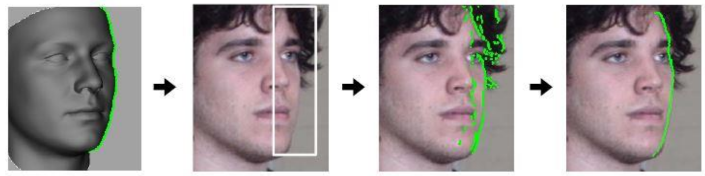
Distinctive features of the face are detected manually or automatically first and co-ordinates are saved. Occluded feature co-ordinates are estimated.

Using Orthographic projection model, a 3D generic shape is aligned to the 2D Image. The 2D image is then back projected to 3D model and frontal face is rendered with textured 3D model



B. Un-Occluded facial texture detection

The 3D generic shape model is aligned with the 2D image and then projected to 2D plane to obtain the model roughly in the pose of the 2D image (a). The contour is then detected. Based on the contour of the 3D model, contour search is limited to certain region as shown in (b)



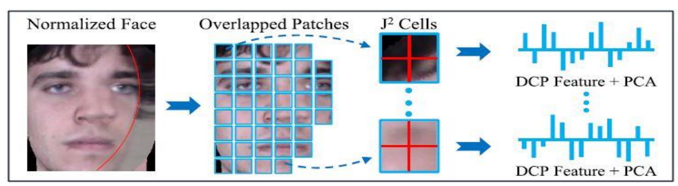
(a) (b) (c) (d)

Edge points are detected with the help of Boundary Detection algorithms(Canny Operator was used in the referance paper) . The facial contour is then obtained by a point sets registration algorithm called Coherent Point Drift (CPD) Briefly, CPD iteratively aligns the facial contour to the edge point set with affine transformations (d).

This step is projected to the 3D model in the Face pose normalization step along with the facial texture. It is then projected to the rendered frontal face.

C. Patch Based Face Representation

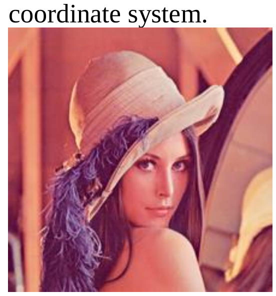
Normalised face is divided into patches. A patch is maintained if in the patch 80% of pixels fall into unoccluded region, else is discarded. All the new Unoccluded patches are divided into JxJ cells.



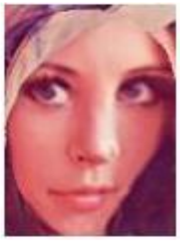
DCP (Dual Cross Patterns) is employed for feature extraction. DCP histogram from JxJ cells form raw feature of the patch. PCA is applied to each patch to project its features into a subspace. After the DCP-PCA, the patches represent the face image.

D. Implementation

Pose estimation - Given a textured 3D model of a face, the synthetic, rendered view of this model is produced by specifying a reference projection matrix CM = AM [RM tM], where AM is the intrinsic matrix, and [RM tM] the extrinsic matrix consisting of rotation matrix RM and translation vector tM. We select rotation and translation to produce a frontal view of the model which serves as our reference (frontalized) coordinate system.



Frontal pose synthesis - An initial frontalized view IF is produced by projecting query facial features back onto the reference coordinate system using the geometry of the 3D model. For every pixel coordinate q0 = (x0; y0)T in the



reference view, we have the 3D location P = (X; Y;Z)T on the surface of the reference which was projected onto q0 by CM. We use the expression

$$p \sim CQ * P$$

to provide an estimate for the location p = (x; y)T in IQ of that same facial feature.

Bi-linear interpolation is used to sample the intensities of the query photo at p. The sampled colour is then assigned to pixel coordinates q0 in the new, frontalized view.

Soft-Symmetry – Due to a large number of pictures where occlusion is a major concern the frontalization algorithm returns unnaturally unsymmetrical frontal faces. This was resolved using conditional soft symmetry where features of the un-occluded side of the face were symmetrically overlaid on the occluded side of the face, making the synthesized image more natural and realistic

Frontalilzed with soft symmetry

Advantage of this representation is that it is general and can be used for arbitary poses. However, since only limited features are used will result in larger normalization error. This is an disadvantage of this representation.

III. Multi Task Pose-Invariant Face Recognition

Since a frontalization implementation has been established we now aim to take multiple pose images of the same subject and get a transformation matrix dictionary for each one of the poses and optimize each matrix through MtFT learning where each transformation matrix learns from the others in the dictionary establishing the correlation between each one of the transformation matrices.

Images of the subject (A student colleague) in 4 different poses were collected as follows

Using these 4 pictures, 4 frontal images were synthesized using the method described in part II

The transformation matrix for each one of these images were added to a transformation dictionary and MtFTL was implemented on this dictionary yielding the optimal transformation matrix for each of the 4 transformations. Soft Symmetry was used to synthesize occluded features.

MTL implicitly increases the sample size and improves the generalization ability for each task and is beneficial for small training data for the task.

MTL provides a principled way for us to model the correlation between poses if we view the learning of feature transformation for each pose as a task.

MtFTL is arguably the first MTL approach that jointly learns feature transformations for different poses and is shown to profit from the latent inter-pose correlations.

IV. VALIDATIONS PROCEDURE

The procedure for validation used was a PCA-based face recognition system called 'Eigenface'. The process followed for face matching is explained as follows:

A dataset of the test images (Optimized frontalized pictures) is created and a database of random frontal face is created which includes the frontal face of the subject. Our objective is for all the images in the test dataset to match with the correct frontal face in the target database

The optimized test dataset and target database is as follows

Test dataset

Target Database

Since the target database is small in size, we can see that we want the matched result as image 2 in the target database for all 4 images in test dataset.

For this we use Eigenface Face recognition.

Eigenface: Eigenface starts off by taking the target database and reshaping all the 2D images into 1D column vectors and puts these 1D column vectors in a row to construct 2D matrix 'T'.

If there are P images in target database each of same size MxN, then T will be a M*NxP 2D matrix.

Principal Component Analysis is used to determine the most discriminating features between images of faces and 3 values are got from matrix T and the target database – m (the mean of the target database (M*Nx1)), Eigenfaces (Eigen vectors of the covariance matrix of the target database (M*Nx(P-1))), and A (Matrix of centered image vectors (M*NxP)).

Now the test dataset is iterated through and Euclidian distance is calculated between the projected test image and the projection of all centred training images. The matched image is shown as the image with the minimum distance in the target database.

V. RESULTS

When Eigenface was run on MATLAB using the two shown test dataset and target database, all four synthesized images matched with the correct frontal face

The framework for synthesis of frontal faces was very inefficient beyond 45 degrees of yaw but showed reasonable results within 45 degrees of yaw.

The proposed frontalization method has also been recorded to have previously cleared the Adience Gender estimation using the LFW Benchmark

Method	Addience-aligned	Adience3D
LBP	0.734 ± 0.007	0.800 ± 0.007
FPLBP	0.726 ± 0.009	0.753 ± 0.010
LBP+FPLBP+Dropout 0.5	0.761 ± 0.009	0.793 ± 0.008

As expected, the error increases as we increase the size of the target database and pose variation from frontal. The same test database which produced 100% accuracy for a target database of 10 images provides 75% accuracy for target database of 10-32 images, 25% for a target database 32-88 images and vary from no matches to 25% match on databases with more than 100 images.

Pose variation beyond 45 degrees causes a significant degradation in match accuracy never yielding more than 42% accuracy on target databases with more than 10 images and decreasing rapidly as the target database increased in number of images.

VI. IMAGE SOURCES

1. Diagrams: [1]
2. Lena: USC-SIPI image database http://sipi.usc.edu/database/database.php?volume=misc&image=12
3. Images for testing: Camera taken images
4. Images for target database: LFW Dataset

Images for algorithm testing: LFW Dataset

OBSERVATIONS AND COMMENTS ON RESULTS

The proposed methodology provided very reasonable results for a small target database and within 45 degrees of yaw. The results degraded for larger databases and drastically decreased for pose variations more than 45 degrees as the synthesized frontal face lacked the element of recognition and only arbitrary facial features could be recognized which could be matched with a very large number of frontal images. The results also degraded on larger datasets as the probability that the Euclidian distance between test image and target image is falsely low (False Positives) also increases with increasing size of the database. Overall, the proposed implementation was functional to a reasonable degree of accuracy for pose variations within 45 degrees and for databases with lesser than 32 images.

REFERENCES

[1] Changxing Ding, *Student Member, IEEE*, Chang Xu, and Dacheng Tao, *Fellow, IEEE "Multi-Task Pose Invariant Face Recognition" (2015)*

[2] Tal Hassner, Shai Harel1, Eran Paz1, Roee Enbar; "Effective Face Frontalization in Unconstrained Images" (2014)

[3] P. N. Belhumeur, J. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. In ECCV (1), pages 45--58, 1996.