Pose Invariant Face Recognition

Jeet Kumar, Aditya Nigam*, Surya Prakash, Phalgungi Gupta
Speaker* : Aditya Nigam

Indian Institute of Technology Kanpur

Tuesday 20th December, 2011
Overview

1. Introduction
   - Problem Statement
   - Contribution

2. Existing Approach
   - Various Face Recognition Approaches

3. Proposed Approach
   - Proposed Approach: LOC Assumption
   - Proposed Approach: Stages

4. Results
   - Results on Various Database

5. Conclusion and Future Work

Speaker: Aditya Nigam (IIT Kanpur)
Socproc 2011, IIT Roorkee
Tuesday 20th December, 2011
Personal Authentication Systems (PAS)

- Biometric based Personal authentication systems are in demand.
- Several biometric traits are studied such as face, iris, palmprint, ear, fingerprint etc.
- Biometrics based PAS:
  - **Authentication Problem** One to One matching and decide using thresholding (Verification).
  - **Identification Problem** One to Many matching and best matching scores and corresponding subjects are reported (Recognition problem)
Several Biometric Traits and Challenges

- **FACE**: Expression, Illumination, Pose, Occlusion, Ageing.
- **IRIS**: Occlusion, Specular reflection, User Co-operation, Difficult to acquire and Very expensive acquisition sensors.
- **FINGERPRINT**: Fail to acquire specially for cultivators and workers, low public acceptance as connected to criminals and Dirty.
- **EAR**: Occlusion, Illumination.
- **PALMPRINT**: Non-uniform illumination, Expensive acquisition and Require too much pressure.
- **NEW TRAITS**: Knuckleprint, Footprint, Vein Patterns etc.

**Figure**: New Biometric Traits
Pose Invariant Face Recognition

Definition

*Pose invariant face recognition* refers to recognising face images in different poses.

- Rotation of face image (rotation in tilt or yaw) induce very large changes in face appearance.
- Recognition rates fall drastically when images from two different poses of same person are matched.

Figure: Sample images

(a) $0^\circ$  (b) $15^\circ$  (c) $-15^\circ$  (d) $30^\circ$  (e) $-30^\circ$
Problem Statement

- Almost all the face databases have frontal faces. Hence, for matching we require the frontal face image or one should have to generate the frontal face features from the non-frontal face image.
- Linear Object Class assumptions ensures that there exists a linear transformation which relates feature vector of two different posed images. \(^1\)
- In this work we have estimated this transformation and generate frontal face features from the features of posed image.

---

Motivation

- Pose tolerant face recognition has wide application. Test subjects can be uncooperative and system can be passive and non-intrusive.
- In a security and surveillance system face images can be in any arbitrary pose. Hence Pose tolerance is essential for non-cooperative subjects and in uncontrolled and less restricted environment.
Contribution

- An efficient pose invariant face recognition system is developed using Principal Component Analysis (PCA) and regression model.
- Major steps of the developed system are:
  - Pre-processing: It crops the face image and removes noise using a 2-level haar wavelet decomposition.
  - Feature Extraction: PCA has been used to extract features from the preprocessed image.
  - Learning Transformation: Transformation between frontal and posed image feature vectors are learnt using an efficient robust regression technique.
  - Matching and Decision: Euclidean distance metric has been used to obtain the matching score between gallery and transformed probe features of the two face images.
Existing Approach

Existing approaches of pose invariant face recognition can mainly be classified into the following three categories:

- **Model based approaches**: Models like Active Appearance Model (AAM)/Active Shape Model (ASM) and Elastic Bunch Graph Matching (EBGM) creates and deforms a generic face model to fit with the input image.

- **Appearance based approaches**: They extract features using Principle Component Analysis (PCA), Gabor filters or directly using the pixel intensity value.

- **3D based approaches**: They derive a 3D morphable face model by transforming the shape and texture of the training images.
Shortcomings

Approaches in the existing literature suffer from some of the following problems:

- They are not automatic and need manual intervention (AAM/ASM).
- Computationally they are very intensive (PCA).
- Least square regression estimate is not accurate due to the presence of statistical phenomena like multicollinearity and heteroscedasticity (Basic Regression Analysis).
Proposed Approach: LOC Assumption

Proposed approach is based on the assumptions of Linear Object Class. LOC assumption states:

1. The transformation relating feature vectors of the given and the desired poses is linear.
2. A given feature vector of a face image can be represented as a linear combination of corresponding training vectors of its pose.
The proposed system consists of the four major blocks:

- Preprocessing
- Feature extraction
- Learning Transformation
- Feature Matching
Stages: Preprocessing (2-level haar wavelet decomposition)

- Noise removal is done using 2-level haar wavelet decomposition on face images to reduce noise and unnecessary details.
- Only the approximation component (low frequency component) is preserved at each level so as to reduce the size of image as well as computation.

(a) Level 1 decomposition  (b) Level 2 decomposition

Figure: Level two Haar decomposition
Stages: Feature Extraction (PCA)

- Principal Components are extracted and used as features. Eigen vectors of the covariance matrix are the basis vectors.
- Dimensionality reduction is done using PCA by selecting only a few basis vectors among the set of basis vectors depending on their significance.
- Significance is decided on the basis of eigen vectors preserving the most of energy, interpreted from their corresponding eigen values.
- Once the significant eigen vectors are selected, the images are projected into the reduced vector space, that is, the vector space formed only by the selected significant eigen vectors.
- Finally projection coefficients forms the feature vector and the reduced vector space forms the feature space.
Stages: Learning Transformation

- Feature vectors of images in vector spaces of each pose are used to learn the linear transformation between the feature vectors of frontal and non-frontal face image feature vectors.

- Feature vectors of corresponding poses are present as column vector in the matrix $A_P$ and $A_F$. Let $U$ is the desired transformation matrix to be learnt.

  \[ A_F = UA_P \]  

- Ordinary Least Square solution is given by:

  \[ U = \left( (A_P^T A_P)^{-1} A_P^T A_F \right) \]  

- OLS solution is not found to be a very good estimate.
Better solution is proposed to handle these challenges and is given by ridge regression.

\[ \hat{\beta} = (A_P^T A_P + \lambda I)^{-1} A_P^T A_F \]  

(3)

\( \lambda \) is called the ridge parameter and is calculated iteratively such that the covariance of the error converge.

\[ \text{error} = A_F - A_P \hat{\beta} \]  

(4)

Error term is refined at each iteration till it reaches below a certain threshold value.

Ridge regression deliberately introduces bias into the estimation of \( \beta \) in order to reduce the variability of the estimate.

Resulting estimate has generally have lower mean squared error than conventional approaches like the OLS estimates.
Stages: Feature Matching

- For any probe image (non-frontal face) feature extraction is done.
- The learned transformation matrix ($\mathbf{U}$) is applied on the feature vector of a non-frontal probe image to obtain the feature vectors of its corresponding front face features.
- The obtained front face feature vector is matched with the front face gallery feature vector of subjects in the database.
- Euclidean distance classifier is used:

$$Matchscore = \sum_{i=1}^{n} (i^{th} \text{ element of probe} - i^{th} \text{ element of gallery})^2$$ (5)
Results on IITK Database

- IITK database is an Indian face database which consists of 109 subjects and each subject has images in poses: $0^\circ$ (front pose) $\pm 15^\circ$ and $\pm 30^\circ$.
- 50 images are used for learning the transform and the remaining 59 are used for testing.

Sample images in IITK database:

(a) $0^\circ$  (b) $15^\circ$  (c) $-15^\circ$  (d) $30^\circ$  (e) $-30^\circ$

Figure: Sample images
Results on IITK Database

Following results are obtained upon matching:

<table>
<thead>
<tr>
<th>Pose Angle</th>
<th>Accuracy(%)</th>
<th>EER(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+15°</td>
<td>89.03</td>
<td>14</td>
</tr>
<tr>
<td>-15°</td>
<td>87.33</td>
<td>13.5</td>
</tr>
<tr>
<td>+30°</td>
<td>82.73</td>
<td>19.3</td>
</tr>
<tr>
<td>-30°</td>
<td>77.22</td>
<td>23.4</td>
</tr>
</tbody>
</table>

**Table**: Results for IITK database
FERET Database

- Consists of face images of persons from different nationality in varying poses and times.
- Contains 261 images in $+15^\circ$ pose, among which 100 are used for learning the transformation matrix and 161 are used for testing.
- Contains 423 images in $-15^\circ$ pose, among which 50 are used for learning the transformation and 373 are used for testing.
- Unlike IITK database which contains only Indian face images, it contains face images of people from different nationality including blacks, whites, asians, caucasians, etc. Hence, results not as good as IITK Database.
Results on FERET Database

<table>
<thead>
<tr>
<th>Pose Angle</th>
<th>Accuracy(%)</th>
<th>EER(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+15°</td>
<td>77.49</td>
<td>23.1</td>
</tr>
<tr>
<td>−15°</td>
<td>79.27</td>
<td>22.6</td>
</tr>
</tbody>
</table>

Table: Results for FERET database

Sample images:

(a) 0° pose  (b) 15° pose  (c) −15° pose

Figure: Sample images in FERET database
CMU PIE Database

- CMU PIE (pose illumination expression) database consists of face images of 63 people under varying face pose, illumination and expression.
- Images having pose angle \( \pm 22.5^\circ \) are used.
- 30 images are used for learning the transform and 33 are used for testing.
- Results are not as good compared to that of IITK database due to presence of people from different nationalities and facial feature.
- A larger pose variation angle, that is, \( \pm 22.5^\circ \) is considered, which further deteriorates the performance.
Results on CMU PIE Database

<table>
<thead>
<tr>
<th>Pose Angle</th>
<th>Accuracy(%)</th>
<th>EER(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>+22.5°</td>
<td>71.78</td>
<td>36.3</td>
</tr>
<tr>
<td>−22.5°</td>
<td>72.73</td>
<td>31.7</td>
</tr>
</tbody>
</table>

**Table:** Results for CMU PIE database

Sample images in CMU PIE database:

(a) 0° pose  (b) 22.5° pose  (c) −22.5° pose

**Figure:** Sample images in CMU PIE database
Comparison With Existing Approach

<table>
<thead>
<tr>
<th>IITK Database</th>
<th>Proposed approach</th>
<th>Existing approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>EER</td>
</tr>
<tr>
<td>+15°</td>
<td>89.03</td>
<td>14</td>
</tr>
<tr>
<td>−15°</td>
<td>87.33</td>
<td>13.5</td>
</tr>
<tr>
<td>+30°</td>
<td>82.73</td>
<td>19.3</td>
</tr>
<tr>
<td>−30°</td>
<td>77.22</td>
<td>23.4</td>
</tr>
<tr>
<td>FERET</td>
<td>+15°</td>
<td>77.49</td>
</tr>
<tr>
<td></td>
<td>−15°</td>
<td>79.27</td>
</tr>
<tr>
<td>CMU PIE</td>
<td>+22.5°</td>
<td>71.78</td>
</tr>
<tr>
<td></td>
<td>−22.5°</td>
<td>72.73</td>
</tr>
</tbody>
</table>

Table: Comparison between existing and proposed approach
Conclusion

- In this paper we have proposed an improved approach for pose invariant face recognition.
- Preprocessed the face images in various poses and then principal component analysis (PCA) is used on the resulting images for feature extraction.
- Transformation matrix is learnt between the two images in pairs by robust regression technique called ridge regression.
- Feature vector is matched with that of its corresponding frontal face image in the database using a euclidean distance based matching algorithm.
- Results are compared with the existing approach and shows improvement.
Thank You