A Comparative Study on Various State of the Art Face Recognition Techniques under Varying Facial Expressions

Steven Fernandes and Josemin Bala
Department Electronics and Communication Engineering, Karunya University, India

Abstract: Through face we can know the emotions and feelings of a person. It can also be used to judge a person’s mental aspect and psychomatic aspects. There are 5 state of the art approaches for recognizing faces under varying facial expressions. These 5 approaches are overlapping Discrete Cosine Transform (DCT), Hierarchical Dimensionality Reduction (HDR), Local and Global combined Computational Features (LGCF), Combined Statistical Moments (CSM), and Score Level Fusion Techniques (SLFT). Matlab code has been developed for all the 5 systems and tested using common set of train and test images. The train and test images are considered from standard public face databases ATT, JAFFE, and FEI. The key contribution of this article is, we have developed and analyzed the 5 state of the art approaches for recognizing faces under varying facial expressions using a common set of train and test images. This evaluation gives us the exact face recognition rates of the 5 systems under varying facial expressions. The face recognition rate of overlap DCT on ATT database was 95% and FEI 99% which was better than HDR, LGCM, CSM and SLFT. But the face recognition rate of CSM on JAFE database, which contains major facial expression variations, was 100% which was better than overlap DCT, HDR, LGCM, and SLFT.

Keywords: Face recognition, DCT, HDR, low-computational features, statistical moments, SLFT.

Received April 30, 2014; accepted June 2, 2014

1. Introduction

There is lot of work done in past years on face recognition and its applications are highly recognized. Some of the applications are access control, surveillance and searching of photographs online. In order to recognize any face, the most difficult task is existence of facial expression variations, because adjoining of dynamic expression in face makes a wide range of variations in recognizing systems. Also, expressions of the face are not communicated verbally. A person shows his emotions in the form of facial expressions. But these will create unclearness in the recognition system. Till today face recognition is one of the widely used research topic in biometric fields and it is rigorously studied in earlier two decades [3, 8, 11, 12, 24, 26, 28, 34, 35, 43, 45, 47, 49]. Recognizing faces under varying facial expressions is still a very challenging task because adjoining of real time expression in a person face causes a wide range of difficulties in recognition systems. Moreover facial expression is a way of nonverbal communication. Facial expression will reveal the sensation or passion of a person and also it can be used to reveal someone’s mental views and psychosomatic aspects. Researcher says that the voice part or spoken words of a content put up only for 7 percent to the effect of the message as a whole, the voice part contributes for 38 percent, but the face will reveal 55% of effect of the said content. This reveals the importance of facial expressions from major point in human communication [5]. From facial expression we can know whether a listener is interested or not interested in the speaker’s words [18]. The face recognition system works in 3 stages. Pre-processing, Feature Extraction, and classification. Pictures or images are normalized before pre-processing so that they can be easily processed in a format in which it can simply process in given condition or scenario [37]. In pre-processing the images are acquired, it is treated as part of it. The faces are acquired by locating them in different backgrounds. Locating accurately the faces in an image is treated as important role in pre-processing in order to get the facial features [27, 29]. In Feature Extraction we extract prominent landmarks surrounding facial components: eyebrows, eyes, nose, and mouth. Facial features are classified: Intransient facial feature and Transient facial feature. If the image containing face is in motion then facial features permanently lie on the face but they can be deformed with change of facial expression [2]. Facial skin textures are affected by transient facial features and they fail to indicate the type of emotions like wrinkles, bulges etc. [13, 21]. During categorization, the transient and in-transient facial features are classified with respect to the desired result. There may be thousands of facial features. Selecting a low dimensional feature subspace from this is key aspect of optimal classification; hence we convert high dimensional input data into low dimensional feature
subspace [4]. “Definite inherent emotions are derived from allied habits”, this idea is introduced by Darwin [6]. He made the assumption that physiognomies are universal across ethnicities and customs which engross basic emotions like happiness, sadness, fear, disgust, surprise and anger. Primarily psychologists examine facial expressions [7]. Using image sequence automatic face recognition is invented by Suwa et al. [36]. Later matured face expressions research is carried out by the efforts of Mase and Pentland in 1990s [19]. Then onwards in 20th century facial expression was a burning topic which was extensively researched under the development of robotics, computer vision and computer graphics [1]. Later Fasel and Luettin [18] and Rothkrantz [25] conducted a brief survey on different contributions to the research in face recognition from 1990 to 2001. But researchers haven’t been able to find perfect solutions.

The current 5 state of the art approaches considered for face recognition under varying facial expressions are overlapping Discrete Cosine Transform (DCT), Hierarchical Dimensionality Reduction (HDR), Local and Global combined Computational Features (LGCF), Combined Statistical Moments (CSM), and Score Level Fusion Techniques (SLFT) [9, 10, 17, 30, 32].

In this article considered all the 5 state of the art approaches for recognizing faces under varying facial expressions. We have used a common set of train and test images from ATT (formerly The ORL Database of Faces) [38], JAFFE (Japanese Female Facial Expression) [40], and FEI (FEI-1 contains frontaliimages_manuallyaligned_part1 and FEI-2 contains frontaliimages_manuallyaligned_part2) [39] face databases. This evaluation gives us the exact face recognition rates of the 5 systems under different facial expressions. The rest of the article is presented as follows: section 2 Overlapping DCT Section 3 presents HDR. Section 4 Local and Global combined Computational Features. Section 5 presents Combined Statistical Moments. Section 6 presents Score Level Fusion Technique. Results are discussed in section 7. Section 8 draws the Conclusion. Section 9 presents Future Work.

2. Overlapping Discrete Cosine Transform

DCT hashing is used for creating index structures for face descriptors [20]. Overlapping-DCT [37] uses local features which allow spatial information about the image to be retained, which is shown in Algorithm 1.

Algorithm 1: Overlapping-DCT algorithm:

Step 1. Start
Step 2. Extract the features of the image by creating blocks of size b x b.
Step 3. Overlap neighbouring blocks by (b-1) x (b-1) pixels.
Step 4. Scan the image from left to right and top to bottom.
Step 5. The number of blocks processed is then: \((M - (b -1)) \times (N - (b -1))\)

Step 6. For each block, the 2-dimensional DCT is then found and \(c_f\) coefficients are retained.

Step 7. The coefficients found for each block are then concatenated to construct the feature vector representing the image.

Step 8. For \(c_f\) retained coefficients and square image, the dimensionality of this feature vector is then:

\[ d = (N - (b -1))^2 \times c_f \]

\(d = \text{feature vector dimensionality}, N \times N = \text{original image resolution}, b \times b = \text{block size}, c_f = \text{number of coefficients retained from each block.} \]

Step 9. End

3. Hierarchical Dimensionality Reduction

High-dimensional datasets has one problem: that is, in many cases all the variables not important for understanding the phenomena of interest. But some mathematically difficult methods can develop predictive models which give high accuracy taking high dimensional data. Still many applications are attempting to reduce the original data dimension before modeling the data.

We have used hierarchical decision tree classifiers [20].

Algorithm 2: HDR algorithm

Step 1. Start
Step 2. Consider a random p-dimensional variable \(x = (x_1, \ldots, x_p)\).
\(x_1 = 1\)-dimensional variable \(x, x_p = p\)-dimensional variable \(x\)
Step 3. Find lower dimensional representation of \(x\), i.e. \(s = (s_1, \ldots, s_p)\) with \(k < p\)
\(s_1 = \text{lower dimensional representation of } x_1, s_k = \text{lower dimensional representation of } x_p\)
Step 4. Hide some components \(s\), this reduces the dimension.
Step 5. Data space is then sampled in order to find a 1-dimensional feature space.
Step 6. Extract the features from the feature space and apply then to hierarchical decision tree classifier.
Step 7. End

4. Local and Global Combined Computational Features

Either entire image features or local features containing image patches are used but most of the face recognition systems. Entire object is generalized with a single vector in global features. But local features make the computation at different points in the image. They are comparatively more robust to occlusion and clutter. We have developed a method which combines global features and local features [22]. LGCF algorithm:

Algorithm 3: LGCF algorithm

Step 1. Start
Step 2. Each lobe of a dissociated multi-poles should be a low-pass filter.
Step 3. Intensity information within the region of the lobe should be statistically estimated.
Step 4. Coefficients of each lobe should be arranged such that the weight of a pixel is inverse proportional to its distance from the lobe centre, we have used Gaussian mask.
Step 5. The sum of all lobe coefficients should be zero, thus the entropy of a single ordinal code is maximized.
Step 6. Large Ordinal feature set contains much redundant information; to remove the redundancy AdaBoost [5] algorithm is used.
Step 7. End

5. Combined Statistical Moments

Among different feature extraction methods proposed for face recognition, statistical moments seem to be so promising [33]. The moments considered are Centralized moments, Normalized moments, Hu invariant moments, and Legendre moments. Centralized moment is a moment of a probability distribution of a random variable about the random variable's mean. The normalized \( n^{th} \) central moment is the \( n^{th} \) central moment divided by \( \sigma^2 \). Hu described two different methods for producing rotation invariant moments, the first method used principal axes and the second is the method of absolute moment invariants. Legendre moments are continuous moments, hence when they are applied to discrete-space image a numerical approximation is involved, this accurate method but it is time consuming. The approach that used for CSM.

Algorithm 4: CSM

Step 1. Start
Step 2. Combine the following moments:
- Centralised moments
- Normalised moments
- Hu invariant moments
- Legendre moments
Step 3. Use each feature vector as not discriminative for identification
Step 4. Use appropriate weights and then perform face recognition
Step 5. End

6. Score Level Fusion Techniques

SLFT [8, 11] is developed using combinatorial approach and Z-Score normalization. The scores generated from Principal Component Analysis (PCA) [14, 44], FisherFaces (FF) [41], Independent Component Analysis (ICA) [23, 31], Fourier Spectra (FS) [48], Singular Value Decomposition (SVD) [16] and Sparse Representation (SR) [46] are fused together. The algorithm used for SLFT:

Algorithm 5: SLFT

Step 1. Start
Step 2. Each algorithm: PCA, FF, ICA, FS, SVD, and SR return a score for each image in the database.
Step 3. The score is calculated as follows:

Scorelist 1: PCA - Euclidean distance
Scorelist 2: FF - Euclidean distance

7. Results and Discussion

We have considered 3 standard public face databases ATT, JAFFE, FEI (FEI-1 and FEI-2) to train and test Overlapping DCT, HDR, LGCF, CSM, and SLFT. The details of train and test images used are tabulated in Table 1.

Face Recognition Rate (FRR) is calculated. For example consider ATT database, it contains 40 subjects, each subject is assigned a class number. Thus for ATT database we have 40 classes. For 40 classes we have 40 folder placed in Train_data folder (s01, s02, s03, …s40). Each class is given a unique class ID, s01 is given class ID 1, s02 is given class ID 2, s03 is given class ID 3 and so on. Now there are total 40 class IDs. Training is performed by taking 2 images per class, 40 x 2 = 80 images are used for training.

Once the system is completely trained then images are picked from Test_data folder, inside Test_folder we will again have 40 folders (s01, s02, s03, …s40), each of these folders will have 10 images. Totally 40 x 10 = 400 images are used for testing. Now when recognition is performed by picking images from Test_data folder, for example if an image is picked from s01 folder, ideally it needs to match to s01 folder of Train_data and it needs to have class ID 1. If it matches to any other images, for example if s01 image is matched to s02, then the class ID would be 2 and this would be considered as a mismatch. Thus FRR is calculated as:

\[
FRR = \frac{\text{Total number of face images recognized accurately}}{\text{Total number of face images available in the database}} \times 100
\]

Figure 1. Recognition using CSM were input image is form JAFFE database.
The results obtained for all the images of ATT, JAFFE, FEI-1, and FEI-2 on all the 5 face recognition approaches can be downloaded from https://drive.google.com/folderview?id=0B0s5kogk-EfAX2FibD4dkxRV0E&usp=sharing FRR is obtained for all the 5 approaches on ATT, JAFFE, FEI-1, and FEI-2 face databases. The results are tabulated in Table 2.

Table 2. FRR calculated for the 5 face recognition approaches.

<table>
<thead>
<tr>
<th>Techniques</th>
<th>ATT Database</th>
<th>JAFFE Database</th>
<th>FEI-1 Database</th>
<th>FEI-2 Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlapping DCT</td>
<td>36/400=95%</td>
<td>193/200=96.50%</td>
<td>198/200=99%</td>
<td>198/200=99%</td>
</tr>
<tr>
<td>HDR - FRR</td>
<td>35/400=87.75%</td>
<td>142/200=71%</td>
<td>196/200=95%</td>
<td>193/200=96.50%</td>
</tr>
<tr>
<td>LGCF-FRR</td>
<td>36/400=91.75%</td>
<td>182/200=91%</td>
<td>198/200=98.50%</td>
<td>198/200=99%</td>
</tr>
<tr>
<td>CSM - FRR</td>
<td>34/400=86%</td>
<td>200/200=100%</td>
<td>196/200=98.50%</td>
<td></td>
</tr>
<tr>
<td>SLFT - FRR</td>
<td>34/400=85.75%</td>
<td>157/200=78.50%</td>
<td>191/200=95.50%</td>
<td>191/200=95.50%</td>
</tr>
</tbody>
</table>

8. Conclusions

In this article an attempt is made to develop and analyze the 5 state of the art approaches for Face Recognition under varying Facial Expression. The approaches considered are: Overlapping DCT, HDR, LGCF, CSM, and SLFT for Face Recognition in the presence of varying facial expressions. All the 5 systems were trained and tested using standard public face databases ATT, JAFFE, FEI (FEI-1 and FEI-2). This evaluation gives us the exact face recognition rates of the 5 systems under varying facial expressions. The face recognition rate of OVERLAP DCT on ATT database was 95% and FEI (FEI-1 and FEI-2) 99% which was better than HDR, LGCM, CSM and SLFT. But the face recognition rate of CSM on JAFÉ database, which measures major facial expression variations, was 100% which was better than overlap DCT, HDR, LGCM, and SLFT. Thus we conclude OVERLAP DCT is the best technique to recognize faces under slight facial expression variations and CSM is the best technique to recognize faces under major facial expression variations.

9. Future work

The 5 state of the art face recognition techniques developed and analyzed for varying facial expression could be used to evaluate the face recognition rate under varying pose and illuminations using standard public face databases: Pointing Head Pose Image Database [41], YALEB Database [42].

References

[8] Fernandes S. and Bala G., “A Comparative Study on ICA and LPP Based Face Recognition Under...
A Comparative Study on Various State of the Art Face Recognition...


Steven Fernandes is the primary author who is a research student at Karunya University, Coimbatore, India. He received his M.Tech in Microelectronics from Manipal Institute of Technology, Manipal, India. His field of interest includes Computer Vision and Machine Learning.

Josemin Bala is the corresponding author who received her Ph.D. from Anna University, Chennai, India. She is a Professor at Karunya University, Coimbatore, India. Her areas of interest are Image Processing, and Wireless Communications.