Face Recognition Using Legendre Moments

Dr.S.Annadurai¹
Professor & Head of CSE & IT
Government College of Technology,
Coimbatore, Tamilnadu, India.
Email: anna_prof@yahoo.co.in

A.Saradha²
Research scholar in CSE
Government College of Technology,
Coimbatore, Tamilnadu, India.
Email: saradha_irtt@yahoo.com

Abstract
The wide range of variations in human face due to
viewpoint, pose, illumination and expression
deteriorate the recognition performance of the existing
Face recognition systems. This paper proposes a new
approach to face recognition problem using Legendre
moments for representing features and nearest
neighbor classifier for classification. The Legendre
moments are orthogonal and scale invariant in its
characteristics and hence it is suitable for representing
the features of the face images. The obtained feature
vectors are transformed using Linear Discriminant
Analysis and stored in the database and are compared
using Nearest neighbor classifier during testing. For
testing the proposed approach, Legendre feature
vector of size 12 is used for the images of ORL
(Olivetty Research Laboratories) database with 40
subjects and each of them having 10 orientations.
Similarly the Hu moments, Discrete Cosine Transforms
(DCT) are also used for feature extraction. The
recognition percentage is compared with the proposed
approach. The recognition percentage of 98.25% is
achieved using Legendre moments which are
comparatively superior than other Face recognition
approaches using central moments (Hu), DCT or other
statistical approaches.

Keywords: Hu moments, orthogonal Legendre
moments, Discrete Cosine Transform (DCT), Linear
Discriminant Analysis (LDA) and nearest neighbor
classifier.

1. Introduction

In recent years, there has been a growing
interest in machine recognition of faces due to
potential commercial applications such as film
processing, law enforcement, person identification,
access control systems etc. Face recognition may
appear to be easy for human and yet computerized face
recognition system cannot achieve a completely
reliable performance. The difficulties arise due to large
variation in illumination, facial appearance, head size,
orientation and change in environment conditions.
Such difficulties make face recognition one of the
fundamental problems in pattern analysis. In general,
Legendre moments [1],[5] and [3] which is invariant to scaling as feature extraction approach. The moments of order 4 which constitute the feature vector of size 12 and is used in the experiments conducted. The proposed approach must be tested with a standard face database in order to prove its efficiency. Hence ORL data base is used which is one of the popular face database used for performance comparison. The extracted features are transformed using Linear Discriminant Analysis (LDA) and nearest neighbor classifier is used for classification. A very high recognition rate of 98.25% is achieved. The results are compared with the recognition performance using features obtained from Hu moments and DCT coefficients.

2. Related Techniques

2.1 Moments for Image Analysis

The region based moments interpretations interpret a normalized gray level image function as a probability density function of a random variable. Properties of this random variable can be described using statistical characteristics called moments [6]. A moment of order (p+q), with image intensity function f(x,y), are defined as

\[
M_{pq} = \int \int x^p y^q f(x,y) dx dy
\]

where the normalizing constant, \( \mu_{pq} \), is given as follows;

\[
\mu_{pq} = \sum_{s} \sum_{t} f(s,t)(s-x_0)^p(t-y_0)^q
\]

The Hu moments which are invariant to translation, rotation and scaling are calculated using the following equation

\[
M_{pq} = \sum_{s} \sum_{t} f(s,t)(s-x_0)^p(t-y_0)^q
\]

The two-dimensional Legendre moments of order (p+q), with image intensity function f(x,y), are defined as

\[
L_{pq} = \frac{(2p+1)(2q+1)}{4} \int \int P_p(x)P_q(y)f(x,y)dx dy
\]

Where Legendre polynomial, \( P_p(x) \), of order p is given by

\[
P_p(x) = \sum_{k=0}^{p} \left( \frac{(-1)^{l-k}}{2^l \cdot l! \cdot (p-k)!} \cdot \frac{(p+k)!}{2} \cdot \frac{\left( \frac{x}{2} \right)^{p+k}}{p-k} \right)
\]

The recurrence relation of Legendre polynomials, \( P_p(x) \), is given as follows;

\[
P_p(x) = \frac{(2p-1) \cdot P_{p-1}(x) - (p-1) \cdot P_{p-2}(x)}{P}
\]

Where \( P_0(x) = 1 \), \( P_1(x) = x \) and \( p > 1 \). Since the region of definition of Legendre polynomials is the interior of \([-1,1]\), a square image of NxN pixels with intensity function f(i,j), 0 ≤ i, j ≤ N-1, is scaled in the region of -1 < x, y < 1. As a result of this, equation (12) can now be expressed in discrete form as

\[
L_{pq} = \lambda_{pq} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_p(x_i)P_q(y_j)f(i,j)
\]

Where the normalizing constant,

\[
\lambda_{pq} = \frac{(2p+1)(2q+1)}{N^2}
\]

\( x_i \) and \( y_j \) denote the normalized pixel coordinates in the range of \([-1,1]\), which are given by

\[
x_i = \frac{2i}{N-1} - 1 \quad \text{and} \quad y_j = \frac{2j}{N-1} - 1
\]
2.3 Discrete Cosine Transform

DCT is a popular technique in transform coding systems and the mean square reconstruction error is less and it has good reconstruction capability. It is also having a good information packing ability. Compared to other input independent transforms it has advantages of packing the most information into the fewest coefficients and hence used for representing the features of the images. A subset of extracted coefficients forms the feature vector and is compared with nearest neighbor classifier. DCT is an orthogonal transform and consists of phase shifted cosine functions. It is calculated using the formula

\[ C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left( \frac{2\pi v x}{2N} \right) \cos \left( \frac{2\pi u y}{2N} \right) \]

For \( u, v = 0, 1, 2, 3, \ldots, N-1 \)

\[ \alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u = 0 \\ \sqrt{\frac{2}{N}} & \text{for } u = 1, 2, \ldots, N-1 \end{cases} \]

2.3. Linear Discriminant Analysis

Discriminant analysis is used to classify cases into one of several known groups on the basis of various characteristics. It can also identify the variables that are good predictors of group membership by using the stepwise method of variable selection.

Available at each step are F statistics for evaluating the usefulness of variables in and out of the equation, Wilk’s lambda or U statistics and F statistics for testing pair wise differences between group means. For the final equation, coefficients are displayed for the classification functions (Fisher’s linear discriminant function) which provide the best discrimination between the groups. The functions are generated from a sample of cases for which group membership is known; the functions can then be applied to new cases with measurements for the predictor variables with unknown group membership.

Linear discriminant analysis easily handles the case where the within-class frequencies are unequal. It maximizes the ratio of between class variance to the within class variance in any particular data set thereby guaranteeing maximal separability.

3. Proposed Face Recognition System

The proposed face recognition system shown in Figure-1 consists of 4 major modules. They are 1) Image normalization 2) Feature extraction 3) Linear discriminant analysis 4) nearest neighbor classifier.

3.1 Image Normalization

The ORL database consists of images of 40 subjects and each subject consists of 10 orientations. Each image is a gray scale image consists of pixel values ranges from 0 to 255. In order to reduce the computational overhead it is normalized by its maximum pixel value.

3.2 Feature Extraction

The next process of face recognition is to extract the invariant features from the normalized images of the database. Three feature extraction approaches are used and the extracted features are subjected to classification. DCT, Hu moments and Legendre moments are used as feature extraction approaches in the experiment. DCT feature vectors are of size 8x8, 16x16 and 24x24. Hu moments contribute a feature vector of size seven and Legendre moments contribute the feature vector of size 12.

![Figure 1 - Block Diagram of Face Recognition System](image-url)
3.3 LDA

The extracted features are invariant features of the images under consideration and are subjected to transformation in order to improve discriminability. These transformed feature vectors are classified with nearest neighbor classifier.

3.4 NN Classifier

The features of the images of the database are extracted, transformed using Linear Discriminant Analysis and stored in the feature database. Similarly during testing the feature vectors of the test image are extracted using DCT, Hu and Legendre moments and transformed using discriminant function. The transformed vector is compared with the respective feature vectors stored in the database by measuring Euclidean distance between the test vector and the feature vectors of the database. A comparative study is performed and the results are tabulated. The experiment is repeated by increasing the number of images of the database from 5 to 40 in steps of 5 and the recognition percentage is calculated and recorded.

4. Experimental Results

The experiments are conducted by extracting the features using the approaches of Hu moments, Legendre moments and DCT. The ORL database is used for experimentation. The ORL database consists of images of 40 subjects of size 92x112 as shown in figure-2 and each subject has 10 orientations. Some orientations of sample image is shown in figure 3. They are resized to 64x64 in order to reduce the computational complexity.

4.1 Experiment 1

The Hu moments are calculated up to third order which results in a feature vector of size 7. Table.1 consists of 10 feature vectors for 10 orientations of the subject shown in figure 3. Feature vectors obtained are subjected to Linear Discriminant Analysis and transformed. The number of subjects used for the analysis is increased from 5-40 in steps of 5 and the recognition percentage is tabulated. Only 40% of recognition percentage is observed.

4.2 Experiment 2

The Legendre moments of order four is considered for the calculation of moments. The extracted feature vectors are subjected to Linear Discriminant Analysis which increases the
Table 1: Feature vectors of sample image using Hu moments

<table>
<thead>
<tr>
<th>IMAGE No.</th>
<th>( \varphi_1 )</th>
<th>( \varphi_2 )</th>
<th>( \varphi_3 )</th>
<th>( \varphi_4 )</th>
<th>( \varphi_5 )</th>
<th>( \varphi_6 )</th>
<th>( \varphi_7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>0.321206</td>
<td>0.000697</td>
<td>0.642788</td>
<td>0.075973</td>
<td>0.016117</td>
<td>-0.00132</td>
<td>0.004703</td>
</tr>
<tr>
<td>2-2</td>
<td>0.31817</td>
<td>0.000378</td>
<td>5.430554</td>
<td>0.290175</td>
<td>0.252781</td>
<td>0.000464</td>
<td>0.262275</td>
</tr>
<tr>
<td>2-3</td>
<td>0.333391</td>
<td>0.000829</td>
<td>3.87237</td>
<td>0.25416</td>
<td>-0.09985</td>
<td>-0.00731</td>
<td>0.231532</td>
</tr>
<tr>
<td>2-4</td>
<td>0.323398</td>
<td>0.000291</td>
<td>2.935851</td>
<td>0.088474</td>
<td>0.020075</td>
<td>0.000133</td>
<td>0.040399</td>
</tr>
<tr>
<td>2-5</td>
<td>0.328096</td>
<td>0.000562</td>
<td>2.368621</td>
<td>0.088565</td>
<td>-0.02353</td>
<td>0.000306</td>
<td>-0.03304</td>
</tr>
<tr>
<td>2-6</td>
<td>0.31771</td>
<td>0.000067</td>
<td>2.95772</td>
<td>0.398794</td>
<td>0.368927</td>
<td>0.00128</td>
<td>-0.22689</td>
</tr>
<tr>
<td>2-7</td>
<td>0.332101</td>
<td>0.001394</td>
<td>9.402906</td>
<td>0.371977</td>
<td>0.473293</td>
<td>-0.00959</td>
<td>0.509858</td>
</tr>
<tr>
<td>2-8</td>
<td>0.338357</td>
<td>0.000946</td>
<td>2.529241</td>
<td>0.010237</td>
<td>-0.00148</td>
<td>-0.0002</td>
<td>0.072834</td>
</tr>
<tr>
<td>2-9</td>
<td>0.330197</td>
<td>0.000017</td>
<td>6.276105</td>
<td>0.631597</td>
<td>1.255381</td>
<td>-0.00252</td>
<td>0.00463</td>
</tr>
<tr>
<td>2-10</td>
<td>0.324126</td>
<td>0.000354</td>
<td>2.406464</td>
<td>0.2113</td>
<td>0.150603</td>
<td>-0.00145</td>
<td>0.0463</td>
</tr>
</tbody>
</table>

Table 2: Feature vector of sample image using Legendre moments of order 4

<table>
<thead>
<tr>
<th>Image No.</th>
<th>( L_{00} )</th>
<th>( L_{02} )</th>
<th>( L_{11} )</th>
<th>( L_{21} )</th>
<th>( L_{12} )</th>
<th>( L_{30} )</th>
<th>( L_{03} )</th>
<th>( L_{40} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-1</td>
<td>0.630146</td>
<td>0.576177</td>
<td>1.016352</td>
<td>-1.02772</td>
<td>-0.97946</td>
<td>-1.52283</td>
<td>-1.50392</td>
<td>2.745453</td>
</tr>
<tr>
<td>2-2</td>
<td>0.591155</td>
<td>0.689645</td>
<td>1.055558</td>
<td>-1.03431</td>
<td>-1.02854</td>
<td>-1.54368</td>
<td>-1.59680</td>
<td>2.757582</td>
</tr>
<tr>
<td>2-3</td>
<td>0.565017</td>
<td>0.631265</td>
<td>1.055558</td>
<td>-1.03285</td>
<td>-1.03285</td>
<td>-1.54368</td>
<td>-1.59680</td>
<td>2.757582</td>
</tr>
<tr>
<td>2-4</td>
<td>0.692927</td>
<td>0.6204</td>
<td>0.957181</td>
<td>-0.87709</td>
<td>-0.86585</td>
<td>-1.47573</td>
<td>-1.53728</td>
<td>2.684474</td>
</tr>
<tr>
<td>2-5</td>
<td>0.375116</td>
<td>0.676499</td>
<td>1.109936</td>
<td>-1.00164</td>
<td>-1.49895</td>
<td>-1.62664</td>
<td>2.741392</td>
<td></td>
</tr>
<tr>
<td>2-6</td>
<td>0.567606</td>
<td>0.70133</td>
<td>0.955957</td>
<td>-0.79248</td>
<td>-0.90876</td>
<td>-1.46515</td>
<td>-1.62683</td>
<td>2.705298</td>
</tr>
<tr>
<td>2-7</td>
<td>0.753509</td>
<td>0.66943</td>
<td>1.16487</td>
<td>-1.18364</td>
<td>-1.10983</td>
<td>-1.65188</td>
<td>-1.61243</td>
<td>2.889758</td>
</tr>
<tr>
<td>2-8</td>
<td>0.66332</td>
<td>0.607345</td>
<td>1.028369</td>
<td>-1.05051</td>
<td>-0.97669</td>
<td>-1.57214</td>
<td>-1.53227</td>
<td>2.798169</td>
</tr>
<tr>
<td>2-9</td>
<td>0.652628</td>
<td>0.749567</td>
<td>1.026864</td>
<td>-0.92732</td>
<td>-0.98328</td>
<td>-1.52831</td>
<td>-1.68223</td>
<td>2.769508</td>
</tr>
<tr>
<td>2-10</td>
<td>0.713316</td>
<td>0.681758</td>
<td>1.101063</td>
<td>-1.11991</td>
<td>-1.10409</td>
<td>-1.63034</td>
<td>-1.63367</td>
<td>2.893186</td>
</tr>
</tbody>
</table>

discriminatory power of feature vectors. The transformed vectors are stored in the database. The number of subjects used for testing is increased from 5 to 40 in steps of 5 and the results are tabulated. A Good recognition percentage of 98.25% is achieved.

4.3 Experiment 3

In this experiment the images are segmented into 8x8. Discrete cosine transform is applied on these segments. For every image segment 2x2 DCT coefficients of DCT matrix is selected which constitutes a feature vector of 16x16 in size. An example 16x16 feature vector of image 1-1.bmp is shown in table 3. For all the images feature vectors of size 16x16 are used for calculating mean and covariance matrices. Testing is done by comparing the covariance matrices of test image with that of images in data base. A recognition percentage of 100% is achieved for 5 subjects. When the number of subjects is increased from 5 to 40 in steps the recognition percentage is slightly degraded and for 40 subjects it is
90%. The same experiment is repeated for different sizes of feature vectors formed by extracting 1x1 and 3x3 coefficients of DCT matrix.

5. Performance Comparison

Using the above experiments the invariant features (Hu and Legendre moments) and DCT coefficients are extracted for the images in the ORL database. The experiments are conducted with varying sizes of data set. The number of subjects considered is increased from 5 to 40 in steps of 5 where each subject contains 10 images. The classification accuracy obtained using both the moments and DCT feature vectors are tabulated in Table.4. It is observed that the Hu moments which are efficient in handling binary images when used for feature representation it is found that it is not an appropriate feature representation method for such a complex face recognition problems. It is also found that the recognition percentage of the face recognition system using orthogonal moments (Legendre moments) is found superior to other feature extraction techniques such as DCT and Hu moments and shown in Figure.5

Table . 4 Performance Analysis

<table>
<thead>
<tr>
<th>No of subjects</th>
<th>% of Recognition</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hu</td>
<td>Legendre</td>
<td>DCT</td>
</tr>
<tr>
<td>5</td>
<td>96%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>10</td>
<td>72%</td>
<td>100%</td>
<td>96%</td>
</tr>
<tr>
<td>15</td>
<td>61.37%</td>
<td>100%</td>
<td>95%</td>
</tr>
<tr>
<td>20</td>
<td>55%</td>
<td>98%</td>
<td>94%</td>
</tr>
<tr>
<td>30</td>
<td>51%</td>
<td>97.36%</td>
<td>93%</td>
</tr>
<tr>
<td>35</td>
<td>50.01%</td>
<td>96.32%</td>
<td>91.03%</td>
</tr>
<tr>
<td>40</td>
<td>46.8%</td>
<td>98.25%</td>
<td>90.01%</td>
</tr>
</tbody>
</table>

Figure.5 Chart for Performance comparison

5.1 Observations from the chart

- Using Legendre moments there is no much degradation in recognition percentage with increasing in the number of images
- Using DCT the recognition percentage is good when the number of images is less and degraded as the number of images increased.

6. Conclusion

In this paper a new method for face recognition is presented. From the experimental results it is observed that the feature representation using DCT provides a maximum recognition rate of 90% and the size of feature vector is also very large. Orthogonal moments that are invariant in nature are capable of representing the images with minimum number of coefficients and the recognition percentage is also superior. Hence it is found very much suitable for feature representation of complex face images which are almost similar in nature with variations size, pose, illumination and orientation within a smaller area. It is also superior over the conventional central moments based recognition system. A good recognition percentage of 98.25% is achieved using Legendre moments. It is also found that it is superior over other statistical approaches like PCA analysis [2] and Line edge map method [7].

7. References