A Performance Study of PCA Based Face Recognition using DCT and DWT Features

Vinay.A¹, Vinay S Shekhar, Ritu Parna J, Tushar Agrawal, K N Balasubramanya Murthy and S Natarajan

PES University, 100 Feet Ring Road, BSK III Stage, Bangalore-560085, Karnataka, India
Email: A.Vinay@pes.edu

Abstract

Face Recognition (FR) has witnessed phenomenal research advances over the past two decades and has been hailed as one of the premier Biometric mechanisms. Feature extractors play a decisive role in dictating the performance of FR systems and a number of effervescent extraction techniques have been proposed over the past few years to increase their efficacy. The effectiveness of these extraction techniques vary distinctly based on the experimental setup, and to that end, Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) are two prominent feature extraction techniques that have been demonstrated to be remarkably effective for FR tasks. In this paper, we perform an exhaustive comparison of the popular variants of the DWT and DCT techniques: DWT-PCA and DCT-PCA, in order to investigate as to which technique yields superior performance.

1 INTRODUCTION AND RELATED WORK

Face Recognition (FR) [1] has received widespread attention from innumerable research communities over the past two decades. It has been established as one of the most exemplary forms of Biometrics, and has proffered a number of cross-domain applications that span from ubiquitous commercial applications such as Access Control and Identity Authentication to critical law-enforcement applications such as Criminal Apprehension through Mug shot matching [3]. Several contemporary applications such as the novel ATM Theft Prevention system [4] and the Driving License Database system, that analyses and tracks driver images to prevent fraud [5] are heavily reliant on FR technology.

Although FR systems have witnessed widespread applicability in constrained applications, their accuracies undergo a sharp decline in unconstrained scenarios [6] (such as extreme variations in expression, pose, illumination, occlusion, scale, resolution and so on) and the general task of FR still poses significant challenges [7]. Furthermore, the ease of availability of a substantial number of facial quality images [3] emphasize the necessity to devise proficient FR mechanisms that are persistently resilient even under intense variations in terms of the aforementioned parameters.

The cumulative performance and reliability of an FR system is dependent upon a variety of factors and among them, the process of feature extraction is of crucial importance [18]. Feature extraction involves the representation of data in a lower dimensional space, which is obtained via a linear/non-linear transformation [14]. A number of state-of-the-art feature extraction techniques have been proposed over the years (a thorough survey of several prevalent techniques can be found in [2]). Recently, the appearance-based approaches have witnessed remarkable success as they function by utilizing the appearance of the face objects and perform the processing of the image as two-dimensional patterns [14]. Among them, Discrete Cosine Transform (DCT) [14][15][16][17] and Discrete Wavelet Transform (DWT) [19][20] have steadily gained traction due to their effectiveness and are commonly utilized in FR tasks. DCT and DWT approaches carry out feature extraction by extracting the most significant coefficients from the training images [21].

Principal Component Analysis (PCA) [11] is widely opted for De-noising and Dimensionality Reduction. The application of PCA to image patches was pioneered in [12] and [10] comprehensively illustrates that it is adequately capable of representing the key-point patches (after they have undergone transformation into a canonical position, orientation and scale) by utilizing it as a local descriptor. We use PCA in our approach as a post-processing step, after feature extraction has been carried out using DCT and DWT, to perform selective key-point selection in order to cut down redundancy and reduce the computational complexity in the classification stage [21] and hence expedite matching performance.

Furthermore, for the face authentication tasks, we make an informed assumption that at least one large face is present in the given complex background (this assumption seems to hold merit as it has exhibited favorable results in [8]). Additionally, we assume that the images considered for the comparison purposes need not be of high resolution unless the set of images that are involved in the experimentations have considerably high likeness [9].

In this paper, we conduct a thorough comparison of the DCT-PCA and DWT-PCA variants on the benchmark ORL [13] face database in order to conclusively determine which approach is more effective. The findings of this study are intended to lend assistance to developers in making informed decisions and hence build robust FR systems.
The rest of the paper is organized as follows: Section 2 outlines the applied DCT-PCA and DWT-PCA techniques, Section 3 details the experimentation results and comparative study and Section 4 proffers a discussion of the proposed approach and outlines future work.

2 APPLIED METHODOLOGY

A. DCT-PCA (Discrete Cosine-Transform-Principal Component Analysis)

Discrete Cosine Transform (DCT) ([14], [15], [16], [17]) is an invertible linear transform that can potently represent a finite sequence of data points as a sum of cosine functions that oscillate at varying frequencies [21]. The given (original) face image undergoes conversion into the frequency domain through the application of two-dimensional DCT. The image can be restored to its original form from the DCT coefficients by utilizing the invert 2D DCT function. The 2-D discrete cosine transform (DCT) [1] is represented in the following manner:

$$X[p,q] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} C_{p,q}[m,n] x[m,n]$$  \hspace{1cm} (1)

Where: \( m = 0, 1, \ldots, M-1, \ n = 0, 1, \ldots, N-1, \)
\( p = 0, 1, \ldots, M-1, \ q = 0, 1, \ldots, N-1, \)

$$C_{p,q}[m,n] = k_p h_q \cos \left( \frac{\pi p (m+1/2)}{M} \right) \cos \left( \frac{\pi q (n+1/2)}{N} \right)$$  \hspace{1cm} (2)

$$k_p = \sqrt{2/M} \ \ \ \text{when} \ p \neq 0, \ k_p = \sqrt{1/M} \ \ \text{when} \ p = 0, \ $$
$$h_q = \sqrt{2/N} \ \ \text{when} \ q \neq 0, \ h_q = \sqrt{1/N} \ \ \text{when} \ p = 0.$$

The characteristic properties proffered by the DCT coefficients in the M x N blocks with respect to the zigzag pattern that are utilized to process the DCT coefficients through JPEG compression are illustrated in Fig.1 [24]. [24] also asserts that even though the total energy is constant throughout the M x N blocks, the distribution of the energy varies and the bulk of the energy is compacted to the low-frequency coefficients. The Discrete Cosine coefficient can then accordingly be represented by \( C(0,0) \) in the forward 2D- DCT equation and since the cosine of zero is one, the equation is simplified to [24]:

$$C(0,0) = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)$$  \hspace{1cm} (3)

The Discrete cosine coefficient that can be found at the upper left corner contains most of the image energy and effectively represents the proportional average of the M x N blocks [16][24]. Accordingly, the remaining ((M x N) – 1) coefficients represent the variation in intensity between the block images, which are called AC coefficients [24]. The DCT can then be performed on the entire image that is obtained after the processing of the input faces through histogram equalization [15][16][24]. This forms the rationale followed by our FR methodology.

In our deliberations, FR is conducted by extracting the important frequency components of the training image that are acquired by the DCT function [21]. The feature vectors of the training images consist of this extracted information. In the subsequent
A Performance Study of PCA Based Face Recognition using DCT and DWT Features

classification stage, the feature vectors thus obtained for the training and test images are meticulously compared using the Euclidean distance [23] similarity measure (section 2.3). A training image that has the minimum distance is employed to perform the classification of the test image.

Now, we suppose that we have an \( r \times c \) image \( f(r,c) \), then the two dimensional \( r \times c \) DCT can be represented as follows [21]:

\[
F(u,v) = \alpha(u) \alpha(v) \sum_{i=0}^{r-1} \sum_{j=0}^{c-1} f(i,j) \cos \left( \frac{(2i+1)\pi}{2r} \right) \cos \left( \frac{(2j+1)\pi}{2c} \right)
\]  \hspace{1cm} (4)

Where
\[
\alpha(u) = \begin{cases} 
\frac{1}{\sqrt{r}} & \text{when } u=0 \\
\frac{1}{\sqrt{2}} & \text{when } u=1,2,3, \ldots r-1
\end{cases}
\]

Further, \( \alpha(v) = \begin{cases} 
\frac{1}{\sqrt{c}} & \text{when } v=0 \\
\frac{1}{\sqrt{2}} & \text{when } v=1,2,3, \ldots c-1
\end{cases} \)

For the DCT -PCA approach, we perform the application of the 2D-DCT function represented in eqn. 1 on the face images in order to render the DCT coefficients. Amidst these coefficients, we only carry out the extraction of the coefficients that represent more relevant information [21]. Subsequently, after the extraction of the relevant frequency components, we apply PCA.

Given a set of data, PCA operates by searching for a new coordinate system, where the principal components (axes) are ordered by the variance that is contained within the training data [22]. In order to accomplish this, the set of face images \( \{x_i\} \) is represented as a matrix \( X \) as depicted below [22]:

\[
X = [x_1 x_2 x_3 \ldots x_M]
\]  \hspace{1cm} (5)

Where the dimension of \( X \) is \( N \times M \) and \( N \) represents the number of pixels in an image [11][22]. Subsequently, we compute the average face and subtract it from each face in \( X \), to obtain \( X' \) [22]:

\[
X' = [(x_1 - \bar{x})(x_2 - \bar{x}) \ldots (x_M - \bar{x})]
\]  \hspace{1cm} (6)

We obtain the principal components of this set by computing the eigenvectors of the covariance matrix \( C \), where [11][22]:

\[
C = \sum_{i=1}^{M} X'X'^T
\]  \hspace{1cm} (7)

The eigenvectors thus computed are utilized as the orthogonal basis in order to represent the training set faces [22]. Consequently, we perform the projection of the training set faces onto the \( k \) eigenvectors that have the highest associated eigenvalues.

The application of PCA on these extracted DCT coefficients yields higher compression rate, along with an improvement in recognition rate [16][21]. Finally, we perform the classification step using the Euclidean distance similarity measure (section A.3).

B. DWT-PCA (Discrete Wavelet-Transform-Principal Component Analysis)

Wavelet Transform is widely employed in image processing and computer vision tasks due to its effectiveness in capturing localized time-frequency information of image extraction [25]. The data is decomposed into different frequency ranges, which permit the effective isolation of the frequency components that are introduced by intrinsic deformations that arise because of expression or other extrinsic factors such as lighting variations etc. into certain distinct sub-bands [25]. Wavelet techniques work by pruning these variable sub-bands, and concentrate on the sub-bands that comprise of the most relevant information that can best represent the given data [19][20][25].

One-Dimensional Continuous Wavelet Transform (CWT) can be represented as follows [25]:

\[
CWT(s,\tau) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} f(t) \psi_{\tau,s}(t) \ dt
\]  \hspace{1cm} (8)

where, \( \psi_{\tau,s}(t) = \frac{1}{\sqrt{s}} \left( \frac{t-\tau}{s} \right) \) denotes the basis function, \( \psi(t) \) is termed the mother wavelet, \( s \) is the scaling parameter and \( \tau \) is called the shift parameter [25]. Discrete Wavelet Transform (DWT) is essentially a sampled version of a CWT. 2-D DWT for a \( m \times n \) image is defined as follows [20][25]:
\[ DWT (j, k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(x) \varphi \left( \frac{x}{2^j} - k \right) \, dx \]  

(9)

Where \( j \) is the power of binary scaling and \( k \) is a constant of the filter.

Discrete Wavelet Transform [19][20] functions by performing the convolution of the target function with the wavelet coefficients that aptly represent the contributions of wavelets in the function at different scales and orientations [19][22]. DWT for an image can be obtained by performing the transformation of the information of a signal into a number that can be sufficiently manipulated, analyzed and reconstructed [19][20][22]. Essentially, the 2-D DWT is computed by successive low-pass and high-pass filtering of an image, along the horizontal and vertical directions [22][25]. 2D DWT decomposes a given image into four sub-bands that represent distinct features: LL (approximate), LH (horizontal), HL (vertical), HH (diagonal). The sub-band LL is approximately at half of the original image; HL consists of the changes of the images (edges) along the vertical direction, while LH is along the horizontal direction; HH consists of high frequency details of the image [20][25]. Figure 2(a) and Fig 2(b) illustrate the different sub-bands in the 2-level and 3-level decomposition of the wavelet respectively [25].

![Fig 2(a) Two level 2-D DWT on an image [25]](image1)

![Fig 2(b) Three Level 2-D DWT on an image [25]](image2)

DWT can be represented as follows [22]:

\[
\begin{align*}
\text{DWT}(x(n)) &= \begin{cases} 
  d_{j,k} = \sum x(n) h_j(n - 2^j k) \\
  a_{j,k} = \sum x(n) g_j(n - 2^j k)
\end{cases}
\end{align*}
\]

(10)

The coefficient \( d_{j,k} \) in eqn.10 is the wavelet function and represents the detail component in the signal \( x(n) \). Accordingly, \( a_{j,k} \) represents the approximation components in the signal and \( h(n) \) and \( g(n) \) functions refer to the coefficients of the high-pass and low-pass filters. Subsequently, \( j \) and \( k \) represent the wavelet scale and translation factors [22].

Subsequently, for the DWT-PCA approach, PCA is applied directly on the DWT extracted coefficients of the face images in order to enforce dimensionality reduction [21]. The steps followed for PCA are along the same lines as those presented in the DCT-PCA approach and involve the extraction of the relevant DWT coefficients prior to the application of PCA. Finally, the classification step is performed using the Euclidean distance similarity measure (section A.3).

C. Similarity Measure

Euclidean Distance is the most commonly employed distance measure [23]. It involves the computation of the root of the square difference between the coordinates of a pair of objects. It is used to calculate the distance between the test image and the image from the database to be identified. We can conclude that the images are similar, provided the distance is small i.e. the distance between the projection of the image and the known projections is computed and consequently, the face image then undergoes classification as one of these faces whose Euclidean distance is minimum [23].

Euclidean distance, \( d_2 \) is computed as follows: \[ d_2(x, y) = \sqrt{\sum_{i=1}^{l}(x_i - y_i)^2} \]  

(11)

3 RESULTS AND COMPARATIVE STUDY

The two approaches have been exhaustively compared on the freely available benchmark ORL [13] database:
The ORL database [13], as illustrated in Fig. 3, comprises of 400 facial images (10 images per individual) with a resolution of 112 x 92 pixels. The images were acquired under a variety of varying parameters such as facial expressions (open / closed eyes, smiling / not smiling) and with/without the presence of occlusions such as glasses. The images were acquired against a dark homogeneous background while ensuring that the subjects are in a frontal, upright position. There is also the presence of affine distortion due to the presence of side-movement of the subjects.

The configurations of the PCs that were utilized in our deliberations is: HP desktop computer with Pentium Dual-Core CPU at 2.6 GHz, 4 GB of memory, and 64-bit Windows 7 Home Premium operating system.

<table>
<thead>
<tr>
<th>No. of Images in Training Set</th>
<th>No. of Images in Test Set</th>
<th>DCT – PCA Accuracy (%)</th>
<th>DWT – PCA Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>84.64</td>
<td>84.64</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>96.25</td>
<td>96.25</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>98.00</td>
<td>98.00</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>97.50</td>
<td>97.50</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>98.33</td>
<td>98.33</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>98.75</td>
<td>98.75</td>
</tr>
</tbody>
</table>

The Experiments have been carried out on 6 configurations of datasets with 6 configurations of the number of PCs. However, from the result perspective, two configurations of PCs have been tabulated. The inferences of the results are:

- The minimum size of the gallery w.r.t each subject (person) and 10 PCs for a good FR system based on DCT-PCA is 4 / 5 and DWT-PCA is 6. Similarly, for 15 PCs, the size still remains 5. This implies that the number of PCs and the training set has an impact on the recognition accuracy.

- The number of PCs does not affect the recognition accuracy of DCT-PCA. The results did not vary when 1, 2, 3, …, 10, …, 15, …, 20 vectors were used during training and testing. On other-hand, DWT-PCA relied on various parameters including PCs i.e. more the number of PCs, better the results.

- The Training set with 6 images / subject w.r.t DCT-PCA yielded less accuracy when compared to that of 5 images / subject.
A thorough accuracy comparison of the the DCT-PCA mechanism against the DWT-PCA can be found in Table.1. Furthermore, the recognition accuracy corresponding to 10 and 15 PCs are furnished in Fig.4 and Fig.5 respectively.

4 DISCUSSION AND FUTURE WORK

We conducted an extensive comparison of the DWT-PCA and DCT-PCA approaches on the benchmark ORL database. We discovered that in certain scenarios, the number of PCs and the size of the training set has an impact on the recognition accuracy of the FR system and after exhaustive experimentations, the minimum size of the gallery for effective FR performance was ascertained to be 4/5 for DCT-PCA and 6 for DWT-PCA. In the case of DCT-PCA, the number of PCs and gallery size had comparatively less influence, as the results did not vary when a varying number of vectors were employed during the training and testing phases. Consequently, it was found that DWT-PCA is more reliant on the aforementioned parameters. Our results establish that the DCT-PCA technique is superior, as it yielded an accuracy increase of 2 to 3 % over DWT-PCA. Due to its robustness over DWT-PCA, DCT-PCA is a more viable choice among the two for FR tasks.

Future work is currently being steered towards the application of different classifiers and dimensionality reduction techniques to DWT along with experimentations utilizing different wavelets to augment the proficiency of FR systems.

REFERENCES


