Faceopedia - High Speed Face Recognition Using DCT RBF Neural Network

Sagar Rai P, Govinda Prakasha*
Department of Computer Science, Centre for PG Studies and Research
St Philomena College, Puttur
Email: gprakash_13@yahoo.com*

Abstract: *The paper titled "FACEOPEDIA"- A High speed face recognition system using DCT RBF Neural Network is designed to identify the face. There are number of algorithm proposed to do the task of face recognition. The face recognition is one of the most trending and new research area where several research activities are taking place. The face is composed of a complex structure; it varies from person to person. There are several applications regarding the face recognition, they are: attendance system, door opening system, authentication system etc. In this project we are using Discrete Cosine Transform (DCT) algorithm along with the Radial Basis Function Neural network. The neural network is needed to identify the complex facial architecture present in the face. The DCT will reduce the dimension and also extracts the essential features of the face.*

Key Words: *Faceopedia, Discrete Cosine Transform, Neural Network, Facial Architecture, Radial Basis Function.*

Introduction

Face recognition is a biometric method of identifying an individual by comparing live capture or digital image data with the stored record for that person. Face recognition is a system using which several real world applications are built to ease the daily activities. Numerous approaches have been proposed for face recognition and considerable successes have been reported. However, it is still a difficult task for a machine to recognize human faces accurately in real-time, especially under variable circumstances such as variations in illumination, pose, facial expression, makeup, etc. The similarity of human faces and the unpredictable variations are the greatest obstacles in face recognition.

Face is a complex multidimensional structure that needs good computing techniques for recognition. The face is our primary focus of attention in social life playing a main role in the identification of individual. We can recognize a number of faces learned throughout our lifespan and identify them at a glance even after years. There may be variations in faces due to aging and distractions like glass, beard or change of hairstyles.

Face recognition is the ability to recognize person by their facial characteristics. Holistic face recognition has attracted more attention since the well-known statistical method, the principal component analysis (PCA). When the face database becomes larger, the time for training and the memory requirement will significantly increase. Moreover, the system based on the PCA should be retrained when new classes are added. As a consequence, it is impractical to apply the PCA in systems with a large database. The discrete cosine transform (DCT) has been employed in face recognition.

Existing System

The face recognition uses the different algorithms to process the different parts that is present in the face. One such algorithm is PCA; it is not efficient method for recognizing the face. The system based on the PCA should be retrained when new classes are added. As a consequence, it is impractical to apply the PCA in systems with a large database. When the face database becomes larger, the time for training and the memory requirement will significantly increase.

Proposed System

The face recognition system is one of the most important areas in modern days. In many organizations, using the face recognition much operation is generally possible. In the proposed system, the DCT algorithm along with the RBF neural network will accomplish the task of identifying the human face more efficiently. The DCT algorithm is mainly used for converting the original image into the reduced dimension and also it will extract the essential features from the human face. The neural network will be used to identify the complex human face architecture. The human face is composed of so many hidden architectures. To identify the human face which is composed of complex architecture is very difficult for a normal computer. The neural network will ease the task of identifying the complex architecture of the human face.

Image Processing System

Image Processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing System includes treating images as two dimensional signals while applying already set signal processing methods to them. It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. The goal of this manipulation can be divided into three categories:

1. Image Processing image in'! Image out
2. Image Analysis image in'! Measusrement out
3. Image Understanding image in'! High-level description out

We will focus on the fundamental concepts of image processing. Space does not permit us to make more than a few introductory remarks about image analysis. Image understanding requires an approach that differs fundamentally from the theme of this book. Further, we will restrict ourselves to two dimensional (2D) Image Processing although most of the concepts and techniques that are to be described can be extended easily to three or more dimensions

Purpose of Image Processing

1. **Visualization** - Observe the objects that are not visible.
2. **Image sharpening and restoration** - To create a better image.
3. **Image retrieval** - Seek for the image of interest.
4. **Measurement of pattern** - Measures various objects in an image.
5. **Image Recognition** - Distinguish the objects in an image.

Types of Image Processing

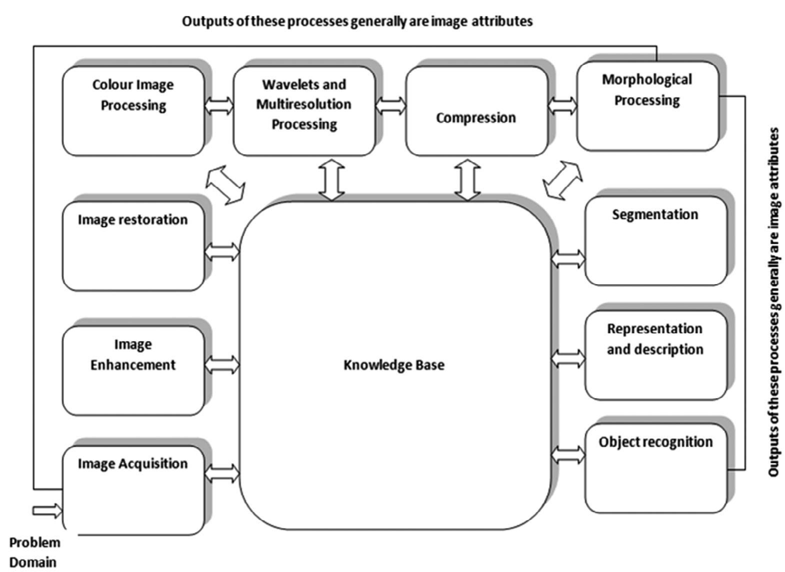
The two types of methods used for Image Processing are Analog and Digital Image Processing.

Analog or Visual Techniques of Image Processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The Image Processing is not just confined to the area that has to be studied but on the knowledge of the analyst. Association is another important tool in Image Processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to Image Processing.

Digital Processing Techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all the types of data have to undergo while using digital technique are pre-processing, enhancement and display, information extraction.

Fundamental Steps in Image Processing:

Fig 1: Digital Image Processing



Hardware and Software Interfaces

The SRS should specify the logical characteristics of each interface between the hardware components and software product.

The hardware interfaces are:

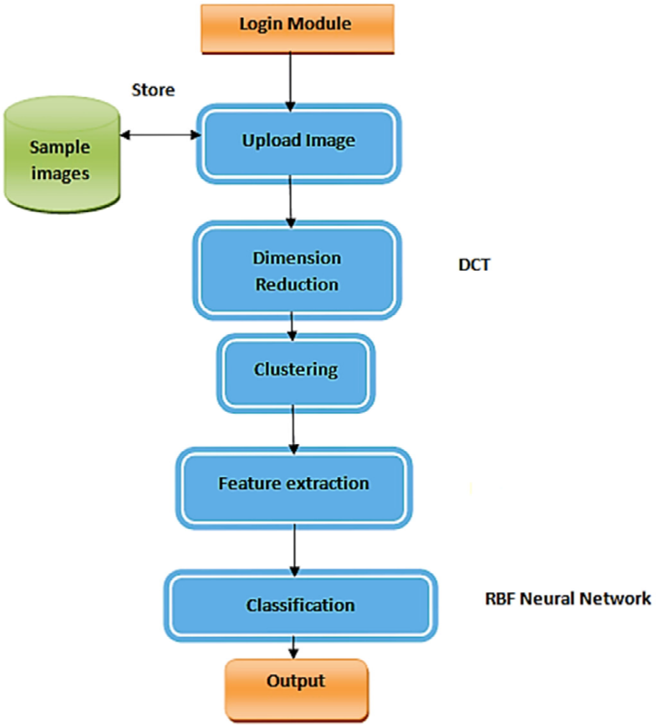
Processor	Intel Dual Core or above
RAM	3GB or above
Hard disk	20GB Hard Disk or above

The software interfaces are:

Operating System	Windows XP or above
Developing Tool	Matlab

Architecture of the System

Fig 2: Architecture of Faceopedia



The High Speed Face Recognition System is based on 5 Modules:

1. Upload image module
2. Dimensionality reduction module
3. Clustering module
4. Feature extraction module
5. Classification module

1. Upload Image Module

The Upload Image Module will allow the user to upload the JPEG image to the system.

2. Dimensionality Reduction Module

Once the JPEG image is uploaded into the system, the dimension of the image is reduced by using the DCT. The DCT has been widely applied to solve numerous problems among the digital signal processing community. For an MXN image, we have an MXN DCT coefficient matrix covering all the spatial frequency components of the image.

Discrete Cosine Transform

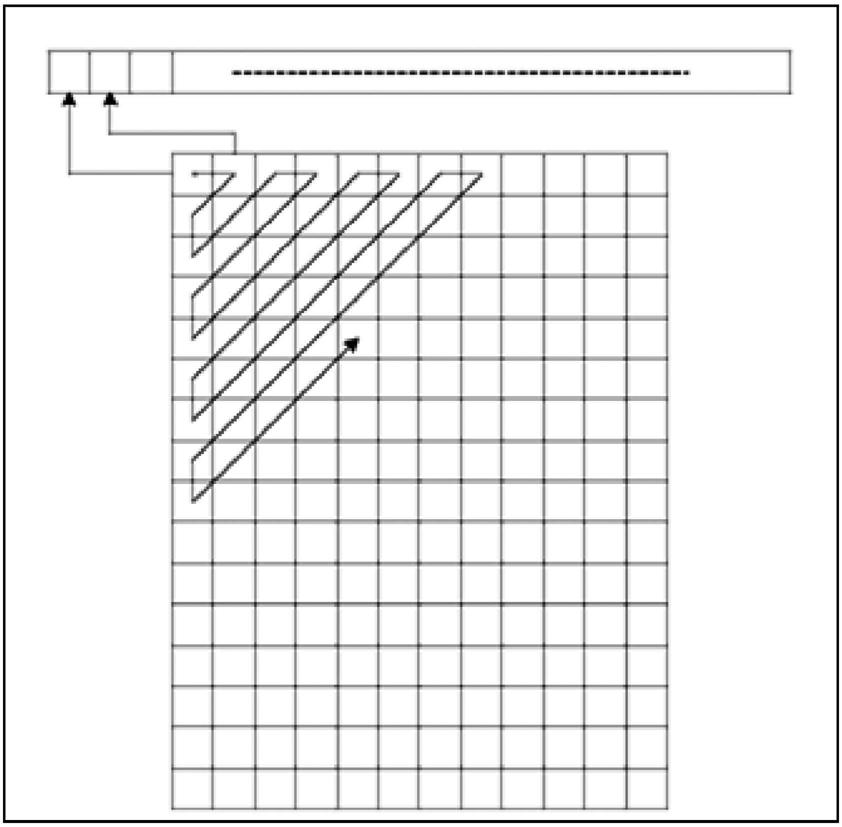
The mathematical theory of linear transforms plays a very important role in the Signal and Image Processing area. They generate a set of coefficients from which it is possible to restore the original samples of the signal. In many situations, a mathematical operation – generally known as a transform – is applied to a signal that is being processed, converting it to the frequency domain. With the signal in the frequency domain, it is processed and, finally, converted back to the original domain. A mathematical transform has an important property: when applied to a signal, i.e., they have the ability to generate decorrelated coefficients, concentrating most of the signal's energy in a reduced number of coefficients.

The Discrete Cosine Transform (DCT) is an invertible linear transform that can express a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. The original signal is converted to the frequency domain by applying the direct DCT transform and it is possible to convert back the transformed signal to the original domain by applying the inverse DCT transform. After the original signal has been transformed, its DCT coefficients reflect the importance of the frequencies that are present in it. The very first coefficient refers to the signal's lowest frequency, known as the DC-coefficient, and usually carries the majority of the relevant (the most representative) information from the original signal. The last coefficient refers to the signal's higher frequencies. These higher frequencies generally represent more detailed or fine information of signal and probably have been caused by noise. The rest of the coefficients (those between the first and the last coefficients) carry different information levels to the original signal.

In the Image Processing field, it is interesting to use a two-dimensional DCT (2D-DCT), because images are intrinsically two-dimensional elements. The standard JPEG, for example, establishes the use a 2D-DCT at the decorrelation step.

In the JPEG image compression standard, original images are initially partitioned into rectangular non overlapping blocks (8X8 blocks) and then the DCT is performed independently on the subimage blocks. In our proposed system, we simply apply the DCT on the entire face image. If the DCT is only applied to the subimage independently, some relationship information between subimages cannot be obtained. However, we can obtain all frequency components of a face image by applying the DCT on the entire face image. In addition, some low-frequency components are only related to the illumination variations which can be discarded. For an image, we have a DCT coefficient matrix covering all the spatial frequency components of the image. The DCT coefficients with large magnitude are mainly located in the upper-left corner of the DCT matrix. We scan the DCT coefficient matrix in a zig-zag manner starting from the upper-left corner and subsequently convert it to a one-dimensional (1-D) vector.

Fig 3: Scheme of Scanning Two-dimensional (2-D) DCT Coefficients to a 1-D Vector



The discrete cosine transform (DCT) is closely related to the discrete Fourier transform. It is a separable linear transformation; that is, the two-dimensional transform is equivalent to a one-dimensional DCT performed along a single dimension followed by a one-dimensional DCT in the other dimension. The definition of the two-dimensional DCT for an input image A and output image B is

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N},$$

$$\begin{cases} 0 \le p \le M-1 \\ 0 \le q \le N-1 \end{cases}$$

where

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \le p \le M-1 \end{cases} \quad \textbf{and} \quad \alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \le q \le N-1 \end{cases}$$

M and N are the row and column size of A, respectively. If you apply the DCT to real data, the result is also real. The DCT tends to concentrate information, making it useful for image compression applications.

3. Clustering Module

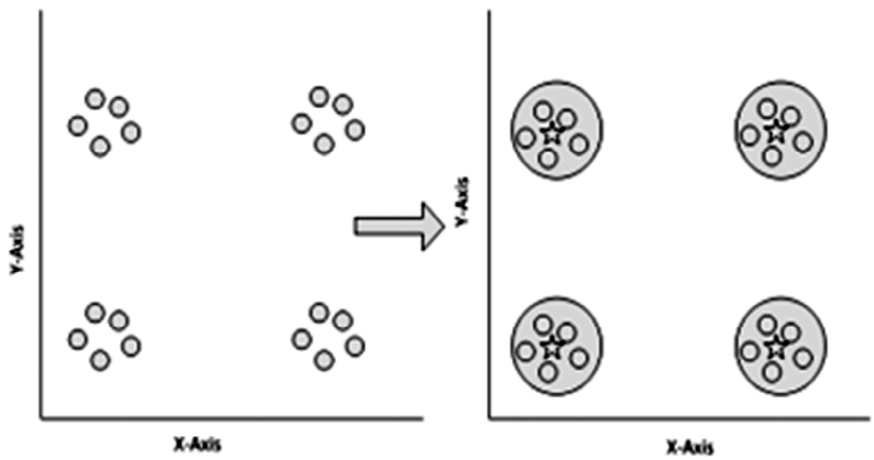
Clustering aims at storing similar or close objects into similar groups in a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise. Figure illustrates to identify four clusters and its centres into which the input data is divided. Two well-known methods of clustering are

3.1 Partitioned clustering

3.2 Hierarchical clustering

In Partition clustering, algorithms find all the clusters simultaneously as a partition of the data and do not impose any sort of hierarchy. In many practical scenarios, there is an inherent hierarchy. The clusters have sub-classes within them, and these subclasses might have their own subclasses. Such classifications are hierarchical and they can be partitioned properly by hierarchical clustering. In partition clustering, the dataset is divided into clusters, such that each cluster has at least one data point and each data point has one cluster.

Fig 3: Clustering of Data



4. Feature Extraction Module

It is used to find a linear projection of the original vectors from a high-dimensional space to an optimal low-dimensional subspace in which the ratio of the between class scatter and the within-class scatter is maximized. After making cluster, the image is passed through high pass filters, wherein image facial parameters can be extracted effectively.

5. Classification Module

Neural Network Based Approaches

Artificial Neural Network (ANN) is a powerful tool for pattern recognition problems. The use of neural networks (NN) in faces has addressed several problems: gender classification, face recognition and classification of facial expressions. One of the earliest demonstrations of NN for face recalls application which was reported in Kohonen's associative map. Using a small set of face images, accurate recall was reported even when input image is very noisy or when portions of the images are missing. In face recognition applications, the RBF neural networks are regarded as a mapping from the feature hyperspace to the classes. Therefore, the number of inputs of RBF neural networks is determined by the dimension of input vectors. The number of outputs is equal to the class number. The hidden neurons are very crucial to the RBF neural networks, which represent the subset of the input data. An artificial neural network is a non-linear and adaptive mathematical module inspired by the working of a human brain. It consists of simple neuron elements operating in parallel and communicating with each other through weighted interconnections.

Model of Neuron

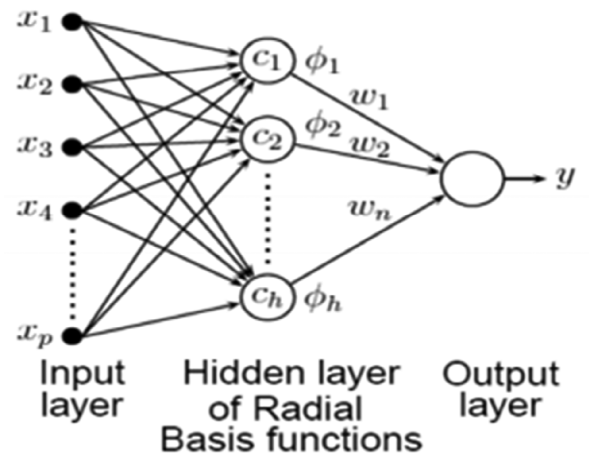
A neuron is an information-processing unit that is fundamental to the operation of a neural network. In this case of artificial neural networks, the strength of the connection between an input and a neuron is defined as the value of the weight. Negative weight values correspond to inhibitory connections, while positive values correspond to excitatory connections. The adder sums up all the inputs modified by their respective weights. Finally, a transfer function controls the amplitude of the output of the neuron. An acceptable range of output is usually between 0 and 1, or -1 and 1 depending on the transfer function selected. Figure 4 shows a typical model of an artificial neuron.

Radial Basis Function Networks (RBFN)

RBNF consists of 3 layers an 1) input layer, 2) a hidden layer, and 3) an output layer.

The hidden units provide a set of functions that constitute an arbitrary basis for the input patterns. Hidden units are known as radial centres and represented by the vectors c_1; c_2; _ _ _ ; c_h transformation from input space to hidden unit space is nonlinear whereas transformation from hidden unit space to output space is linear dimension of each centre for a p input network is px_1

Fig 4: Typical Model of an Artificial Neuron



1. Input Layer

There is one neuron in the input layer for each predictor variable. In the case of categorical variables, N-1 neurons are used where N is the number of categories. The input neurons (or processing before the input layer) standardizes the range of the values by subtracting the median and dividing by the inter quartile range. The input neurons then feed the values to each of the neurons in the hidden layer.

2. Hidden Layer

This layer has a variable number of neurons (the optimal number is determined by the training process). Each neuron consists of a radial basis function centred on a point with as many dimensions as there are predictor variables. The spread (radius) of the RBF function may be different for each dimension. The centres and spreads are determined by the training process. When presented with the x vector of input values from the input layer, a hidden neuron computes the Euclidean distance of the test case from the neurons centre point and then applies the RBF kernel function to this distance using the spread values.

3. Output Layer

The value coming out of a neuron in the hidden layer is multiplied by a weight associated with the neuron (w_1, w_2, ...,w_n in this figure) and passed to the summation which adds up the weighted values and presents this sum as the output of the network. The bias value of 1.0 (not shown in the figure) that is multiplied by a weight w_0 and fed into the summation layer. For classification problems, there is one output (and a separate set of weights and summation unit) for each target category. The value output for a category is the probability that the case being evaluated has that category.

The radial basis function in the hidden layer produces a significant non-zero response only when the input falls within a small localized region of the input space. Each hidden unit has its own receptive field in input space. An input vector xi which lies in the receptive field for centre c_j, would activate c_j and by proper choice of weights the target output is obtained. The output is given as

$$y = \sum_{j=1}^{h} \varphi_j \omega_j, \ \varphi_j = \varphi(\|x - c_j\|)$$

The different radial functions are given as follows

Gaussian Radial function $\varphi(z) = e^{\frac{-z^2}{2\sigma^2}}$

Thin Plate spline $\varphi(z) = z^2 log z$

Quadratic $\varphi(z) = (z^2 + r^2)^{1/2}$

Inverse quadratic $\varphi(z) = 1/(z^2 + r^2)^{1/2}$

Here $z = \|x - c_j\|$

The most popular radial function is Gaussian activation function

Learning in RBFN

Training of RBFN requires optimal selection of the parameters vectors c_i and w_i, i = 1; _ _ _h. Both layers are optimized using different techniques and in different time scales. Following techniques are used to update the weights and centres of a RBFN.

1. Pseudo-Inverse Technique
2. Gradient Descent Learning
3. Hybrid Learning

Training RBF Networks

The following parameters are determined by the training process

1. The number of neurons in the hidden layer.
2. The coordinates of the centre of each hidden-layer RBF function.
3. The radius (spread) of each RBF function in each dimension.
4. The weights applied to the RBF function outputs as they are passed to the summation layer.

Conclusion

This paper presents a high-speed face recognition method based on the techniques of DCT and RBF neural networks. Facial features are first extracted by the DCT which greatly reduces dimensionality of the original face image as well as maintains the main facial features. Compared with the well-known PCA approach, the DCT has the advantages of data independency and fast computational speed. Besides, we have explored another property of DCT. It turns out that by simply discarding the first DCT coefficient, the proposed system is robust against uniform brightness variations of images. Furthermore, by discarding the first few low-frequency DCT coefficients, the effect of no uniform illumination can be alleviated. The architecture and parameters of RBF neural networks are determined according to the distribution properties of the training samples. Simulation results on three benchmark face databases show that our system achieves high training and recognition speed, as well as high recognition rate. More importantly, it is insensitive to illumination variations.

References

http://www.face-rec.org/

https://www.ijircce.com/upload/2015/december/82_Face1.pdf

Rabia Jafari and Hamid Arabnia, R. (2009). A Survey of Face Recognition Techniques. *Journal of Information Processing System.*

Rafael Gonzalez, C., Richard Woods, E. *Digital image processing* (3rd edition).