***Abstract:*** Face recognition is an active field of research with varied applications. This paper discusses about the implementation of  a real time face recognition using invariant image moments. Feature extraction and classification are two key steps for face recognition. We compared three automated methods for face recognition using invariant image moments for feature extraction. For classification, SVM (Support Vector Machine) and k-Nearest Neighbors were used. The experiments were implemented on three face databases, The ATT Face database [1] and the Georgia Tech Database [2] and MIT CBCL database implemented the model for a real-time face recognition system.

***INTRODUCTION***

Face recognition is prominent area of research; there are multitude of algorithms and techniques available in literature to perform face recognition. Feature extraction and classification are two key steps in face recognition. Feature extraction can be defined as the procedure of extracting relevant information from a face image. This can be done in two steps. One is face detection and the other is extracting the features from the detected face. For face-detection Viola-Jones algorithm is used. For feature extraction, Moment functions are used.

The face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. There are three main contributions of this Viola Jones object detection framework. These contributions are evaluated to obtain the feature from the image using the process of integral image followed by cascading classifier and adaboost training to eliminate the weak features.

Moment functions of the two-dimensional image intensity distribution are used in a variety of applications like visual pattern recognition, object classification, template matching, edge detection, pose estimation, robotic vision and data compression. Image moments that are invariant with respect to the transformations of scale, translation and rotation find applications in areas such as pattern recognition, object identification and template matching. Orthogonal moments have additional properties of being robust in the presence of image noise and have a near zero redundancy measure in a feature set. The Legendre moments are invariant to scaling. In general Face Recognition (FR) system suffers with inefficient feature extraction and classification. The success of FR system depends upon either on the feature extraction approach or on the classifier design. The Legendre moments are scale invariant and orthogonal which are found to be suitable for representing the holistic features of the face images in the database. The obtained features from the input images are recognized using K-Nearest-Neighbour (kNN) classifier using the Manhattan distance and Multiclass Support Vector Machine (SVM) classifier.

The outline of the paper is as follows: Section 2 feature extraction In Section 3 classification.  Section 4 contains experimental results and real-time implementation Section 5 concludes the paper.

***IMAGE MOMENTS***

***Legendre Moments:***

Legendre Moments (LM) are continuous and orthogonal moments, they can be used to represent an image with minimum amount of information redundancy. The Legendre moments are scale invariant and orthogonal which are found to be suitable for representing the holistic features of the face images in the database. The obtained features from the input images are used for recognition purpose.

Legendre moments of order g = ( p + q) for an image with intensity function f (x, y) are defined as

where, Pp (x) is the pth order Legendre polynomial defined as

Where *x* ∈ [−1*,* 1], and the Legendre polynomial *Pp(x)* obeys the following recursive relation:

With *P*0*(x)* = 1, *P*1*(x)* = *x* and *p>*1. The set of Legendre polynomials {*Pp(x)*} forms a complete orthogonal basis set on the interval [−1*,* 1]. The orthogonality property is defined as

A digital image of size *M* ×*N* is an array of pixels. Centres of these pixels are the points, where the image intensity function is defined only for this discrete set of points ∈ [−1*,* 1]X[−1*,* 1]. ,are sampling intervals in the *x*- and *y*-directions, respectively. In the literature of digital image processing, the intervals and are fixed at constant values, and, respectively. Therefore, the points will be defined as follows:

With *i*=1*,* 2*,* 3*. . . M* and *j* =1*,* 2*,* 3*. . . N*. For the discretespace version of the image, Eq. (3.6.1) is usually approximated by

Eq. (3.6.5) is so-called direct method for Legendre moments computations, which is the approximated version using zeroth-order approximation (ZOA). As indicated by Liao and Pawlak, Eq. (3.6.5) is not a very accurate approximation of Eq. (3.6.1). To improve the accuracy, they propose to use the following.

Approximated form:

where

Liao and Pawlak propose (AESR) method to evaluate the double integral defined by Eq. (3.6.6), and then they use it to calculate the Legendre moments defined by Eq. (3.6.7).

***EXACT LEGENDRE MOMENTS***

Many algorithms are developed for the computation of LM [6, 7, 8], but these methods focus mainly on 2D geometric moments. When they are applied to a digital image, a numerical approximation is necessary. The difficulty in the use of Legendre moments is due to their high computational complexity, especially when a higher order of moments is used. Error due to approximation increases as the order of the moment increases. An accurate method for computing the Exact Legendre Moments (ELM) proposed by Hosney [9] is as follows

The set of Legendre polynomials {Pp(x)} forms a complete orthogonal basis set on the interval [-1, 1]. A digital image of size NxN is an array of pixels. Centres of these pixels are the points (xi, yj). To improve accuracy, it is proposed to use the following approximated form. One of the special results involving Legendre polynomial is that,

Where *p* =1. For simplicity, upper and lower limits of the integration in Eq. (3.6.8) will be expressed as follows:

Similarly,

Using Equations. (3.6.6), (3.6.7), and (3.6.8), the integral parts will be written as follows:

The set of Legendre moment can thus be computed exactly [1] by

Where

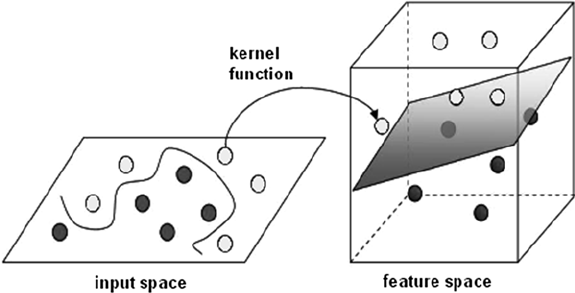
Eq. (3.6.17) is valid only for *p*≥1, and *q*≥1

The moment kernel of exact 2D Legendre moments is defined by Eq. (3.6.17). This kernel is independent of the image. Therefore, this kernel can be pre-computed, stored, recalled whenever it is needed to avoid repetitive computation

***3. CLASSIFICATION***

***3.1 SUPPORT VECTOR MACHINE***

Support Vector Machines (SVMs) are supervised learning methods [14, 16] used for image classification. It views the given image database as two sets of vectors in an ‘ n ’ dimensional space and constructs a separating hyper plane that maximizes the margin between the images relevant to query and the images not relevant to the query. SVM is a kernel method and the kernel function used in SVM is very crucial in determining the performance. The aim of SVM classification method is to find an optimal hyper plane separating relevant and irrelevant vectors by maximizing the size of the margin (between both classes). This is due to the fact if the separating plane has the largest distance to the nearest training data points of any class, it lowers the generalization error of the overall classifier.

******

***Fig1. SVM Kernel function mapping from a non-linear pattern to a linear plane***

***3.2 K-NEAREST NEIGHBORS***

The k-nearest neighbor algorithm [12,13] is a method for classifying objects based on closest training examples in the feature space. It is among the simplest of all machine learning algorithms. Training process for this algorithm only consists of storing feature vectors and labels of the training images. In the classification process, the unlabelled query point is simply assigned to the label of its k nearest neighbors. Typically the object is classified based on the labels of its k nearest neighbors by majority vote. If k=1, the object is simply classified as the class of the object nearest to it. When there are only two classes, k must be an odd integer. However, there can still be ties when k is an odd integer when performing multiclass classification. After we convert each image to a vector of fixed-length with real numbers, the distance function for KNN which is Manhattan distance.

d(x,y) =

Ref: [12] Bremner D, Demaine E, Erickson J, Iacono J, Langerman S, Morin P, Toussaint G, Outputsensitive algorithms for computing nearestneighbor decision boundaries, Discrete and Computational Geometry, 2005, pp. 593–604. [13] Cover, T., Hart, P. , Nearest-neighbor pattern classification, Information Theory, IEEE Transactions on, Jan. 1967, pp. 21-27.

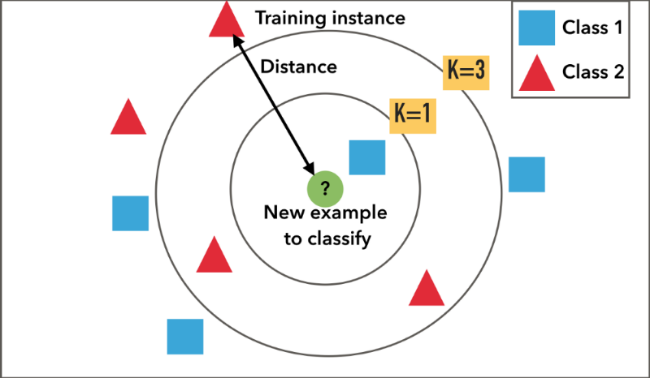


Fig2: Example of k-NN classification. The test sample (inside circle) should be classified either to the first class of blue squares or to the second class of red triangles. If k = 3 (outside circle) it is assigned to the second class because there are 2 triangles and only 1 square inside the inner circle. If, for example k = 5 it is assigned to the first class (3 squares vs. 2 triangles outside the outer circle).

***4. EXPERIMENTAL RESULTS***

***4.1***

Our experiments were performed on two face databases,

The AT&T database consists of 40 subjects each having 10 different variations comprising of 400 images. These images are captured with the dark background under different illumination conditions with varying facial expressions. The image resolution is 92x112 pixels. The AT&T database contains images with very small changes in orientation of images for each subject involved which includes frontal and slight tilt of the head and is in gray-scale.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

The Georgia Tech database contains total 750 images, 15 images each for 50 individuals. The images are in JPEG format and all are of different sizes. The images in the database show variations in illumination conditions, facial expression, and appearance. The images show frontal and/or tilted faces with different facial expressions, lighting conditions and scale. each image for each subject is oriented in a different angle compared to the other.

The MIT-CBCL face recognition database contains face images of 10 subjects.The training set contains High resolution pictures, including frontal, half-profile and profile view of the person. The test set consists of 200 images per subject. The illumination, pose (up to about 30 degrees of rotation in depth) and the background is varied and the images are in 19x19 grayscale PGM format images.

We have mainly used MATLAB version 16.0 for the experimentation purpose.

REAL TIME FACE DATABASE CREATION:

In this process, we have taken 5 subjects and captured 15 facial gestures for each subject thereby comprising of 75 images in the database. These images are captured under same background with different facial expressions under bright illumination conditions. The images thus taken were performed segmentation using Viola-Jones algorithm and later the images were resized to remove the background effects on the image. Then these images were trained to create a real-time face database. After creation of the database we have tested the working in two conditions by taking an input image from the PC and second by taking the live input using the camera and verified the precision in recognition process.



|  |  |  |
| --- | --- | --- |
|  | NN CLASSIFIER | SVM CLASSIFIER |
| REAL TIME FACE DATABASE EFFICIENCY | 74/75  (98.66%) | 75/75  (100%) |

For NN Classifier efficiency:

|  |  |  |  |
| --- | --- | --- | --- |
|  | AAT&T database | Georgia Tech database | MIT CBCL face database |
| Legendre Moments ZOA | 98.5 | 84.26 | 98.5 |
| Exact Legendre Moments | 98.75 | 82.8 | 100 |

For SVM Classifier efficiency:

|  |  |  |  |
| --- | --- | --- | --- |
|  | AAT&T database | Georgia Tech database | MIT CBCL face database |
| Legendre Moments ZOA | 98.5 | 87.56 | 99 |
| Exact Legendre Moments | 99 | 89.75 | 100 |

Using Viola Jones algorithm faces are detected and extracted.

All the detected faces from the video are compared one by one with the face database. If any face is recognized based on the moment features extracted from the input face with that of the faces in the database and then the matched image and name of the image is displayed.