

Comparison of PCA and Image Moments for Hand Gesture Recognition

Surve Pranjali

Department of Electronics (Digital Systems)
AVCOE, Sangamner, India,
Affiliated University of Pune

Ubale V. S

Department of Electronics (Digital Systems)
AVCOE, Sangamner, India,
Affiliated University of Pune

Abstract--This paper presents four simple and efficient methods to implement gesture recognition of hand namely Gradient, Principal Components Analysis, Subtraction and Rotation Invariant. Four different hand gesture images are saved in image database which is created by using web camera. This database is stored in the system for comparison with the input images. Before taking too many images of hand the database is pre-processed in Hand Gesture Recognition system. Pre-processing is very much needed to be done in hand gesture recognition system. To minimize noise by filtering and segmentation are done. In this method the images which are taken by camera which are stored as database in system are compared with the input images to find the best match by extracting features from images

Index terms—Principal Component Analysis, Image Processing, Rotation Invariant

I. INTRODUCTION

This paper is based on the implementation and study of hand gesture recognition systems. In this paper, four different types of hand-gestures are identified- fist, one, two palm. We will study different methods that will give best match results of recognition. In this for the recognition purpose real time camera is used, so the process becomes fast. This system can be used anywhere such as home, office because in this USB camera is used. MATLAB 7.1 software is used in this system

Database is created for all the four hand gestures using ten different images. The images are labelled using integer numbers starting from 1. Following methods were studied and implemented

- **Gradient Method:** In this we have to implement gradient magnitude calculation. Then we have to apply threshold in the gradients
- **Rotation Invariant:** This method is used to provide scale, orientation and position invariant representation of the object

- **Subtraction Method:** This method involves converting all the images into black and white. It involves a simple subtraction between two images
- **Principal Component Analysis:** It is a useful statistical technique developed for regression, reduction of dimensionality, and noise reduction

Main aim of hand gesture recognition system is to identify hand gestures and to classify them accurately. Performance of most algorithms is reported on identification task since characterization of task holds on identification. To successfully meet verification, for hand gesture recognition systems, it is needed to create testing and training databases that specifically address these applications. The process will start with image processing techniques such as, segmentation, noise removal (low-level) feature extraction.

II. SYSTEM OVERVIEW

Any gesture recognition system works according to following steps first acquiring the input image from camera, filtering, segmentation, feature extraction and classification. To control robot or to convey meaningful information we have to create human computer interaction to recognize gesture for physically disable people. Our aim is to make computer understand any human gesture to control various applications

A. Extraction method and image preprocessing

A very good segmentation is required to choose threshold of gray level to extract hand from background. i.e. no part of object should be present in the background and vice versa. With help of segmentation we can separate regions from boundaries. Various segmentation algorithms are there for various applications

Depending upon type of hand gesture, for dynamic hand gesture, we have to locate and track hand gesture. For static gesture (posture) first input image is segmented. A bounding box is created to locate a hand firstly depending on the skin

color and secondly, we need to track the hand. If we want to track hand from a video, the video is divided into frames and all the frames are processed individually considering it as a static posture

B. Features Extraction

A very important step in gesture recognition is feature extraction. After extraction of features it is given as an input to the classifier. If the segmentation is done perfectly it produces features that play important role in recognition process. In feature extraction first we have to find edge of the segmented and filtered image. We come to know about boundaries of different objects due to edge. Edge can be said as sudden change in the intensity from one pixel to other pixel. Edge detection leads to reduction in some amount of data but same shape is maintained.

Depending upon the application there are many different ways of feature extraction of the segmented image.

C. Gestures Classification

The task of assigning a feature vector or a set features to predefined classes in order to recognize any gesture is called as classification. After analysis and modelling of the input hand image is done, to recognize the gesture, a gesture classification method. Many classification methods have been proposed and tested successfully in various recognition systems. Process of recognition is affected with the proper selection of suitable classification algorithm and features parameters.

Gestures can be meaningful and expressive body motions involving physical movements of the body, face, arms, head, fingers, hands with the intent of:

- 1) *Interacting with the environment.*
- 2) *Interaction with physically disable people.*
- 3) *Conveying meaningful information*

A class is a set of reference features that were obtained during the training phase using a set of training images. For the features extracted in the previous phase finding the best matching reference features is classification. For the information to subsequently reconstructed and to be transmitted elsewhere by the receiver a gesture can be perceived by the environment as a technique of compression. Gesture recognition has wide-ranging applications as following:

- 1) *Visualization*
- 2) *Computer games*
- 3) *Man-machine interface*
- 4) *3D animation*
- 5) *Control of mechanical systems*

Gestures can be static or dynamic, the different types of gestures are:

A. Head and face gestures

- 1) *Nodding or shaking the head*
- 2) *Raising the eyebrows*

B. Hand and arm gestures

- 1) *Pointing gestures*
- 2) *Isolated signs*

C. Body gestures in which there is involvement of full body motion as in

- 1) *Formedical rehabilitation and athletic training, recognizing human gaits.*
- 2) *Tracking movements of two people interacting outdoors;*
- 3) *Analysis of movements of a dancer for generating graphics and matching music.*

D. Hand and Arm gestures:

- 1) *Recognition of sign languages, hand poses, and entertainment applications*
- 2) *Head and Face gestures: some examples are:*
 - a) *Raising the eyebrows;*
 - b) *Winking;*
 - c) *Nodding or shaking head*
 - d) *Opening the mouth to speak;*
 - e) *Direction of eye gaze,*
 - f) *Looking surprised; and*
 - g) *Flaring the nostrils, happiness, disgust, fear, anger, sadness, contempt, etc.*

III. RELATED WORKS

Camera is used to pick the gravity center of the points and hand with maximal distances, in a multisystem, from the center gives the locations of tips of the finger, which are then utilized to obtain a skeleton image. In a special camera that supplies dept information was used to identify hand gestures. Other computer vision methods used for hand gesture recognition include particle filters, orientation histogram, neural networks, Fourier descriptors, principal component analysis, specialized mappings architecture. The classification is performed in the curvature space in visual hand recognition system. It involves to find the boundary contours of the hand and it is robust in rotation, scale, translation, it is extremely demanding.

We classify the extracted feature sets using a multiclass SVM that was trained in a setup phase and is trained to differ between predefined gestures. The system works in non-uniform backgrounds and in moderately noisy environments and is suitable for both offline classification and real-time. We

focus on hand gesture classification using inputs from low-resolution digital cameras and classify the acquired images using features that are extracted by non-computationally intensive image processing techniques.

III. GRADIENT METHOD

In this first of all, the gradient magnitude calculation has to be implemented. To define in the picture where the biggest gradient magnitudes are. To apply a threshold in the gradients will be easy. In order to keep the interesting one and to cut down all noise present in the background. For this magnitude is calculated with the formula:

$$\text{Mag} = dx^2 + dy^2$$

To erase the background defects is the goal of this filter. To have a homogeneous picture and to blur the image we had to realize a Gaussian filter. Better results can be obtained in gradient magnitude. To avoid noise it is important to have a uniform background. Lower levels gradients have to be erased since we created a gradient magnitude. Because of this all the noise will get cut and regularize the background. With the Gaussian filter this part will be complimentary. The big defects will be blurred by the Gaussian filter and the lowest magnitudes will be cut by threshold. Then the noise will be reduced. The Euclidian distance of the different images is to be calculated and it is the distance between different vectors. Now different histograms will be compared. This is the final step. Different gestures can be recognized now. Special mathematical background is not required by this method.

IV. SUBTRACTION METHOD

In this we have to convert all the images into black and white including test images. It is a very simple method to implement. This method is not very efficient one because the result generated may be highly inaccurate. It will be done by doing selection of threshold value for the pixels i.e., suppose p be the pixel intensity value and T be the selected threshold value then replace $p=255$ if $p \geq T$ and $p=0$ if $p < T$. Euclidean distance can be calculated by subtraction of each pixel in the test image with the related pixel each of the images in the database. The smallest value of d' will indicate the nearest match. Here d' is defined as

$$d = \sqrt{\sum_{x=0}^{\text{width}} \sum_{y=0}^{\text{height}} (f1(x,y) - f2(x,y))^2}$$

VI. PRINCIPAL COMPONENT ANALYSIS (PCA) METHOD

Some mathematical background is required, for the analysis of principal component analysis for hand gesture recognition. It is called: Eigenfaces or PCA. PCA finds applications in various fields such as face recognition, image compression. High dimensional data can be found with the help of this common technique. We will study few concepts of mathematics that are used in PCA:

a. Mathematical definitions:

1) Standard deviation s ,

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}}$$

2) Standard Deviation:

For the notation, the symbol X will be used so that entire sample can be referred and the symbol X_i will be used to indicate a particular data of the sample. We often use population samples so that measurements can be realized.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

3) Covariance

Covariance can be expressed as:

$$\text{var}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})}{(n-1)}$$

4) Variance

In a set another measure of the spread out of data is variance. It is same as the standard deviation.

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}$$

5) Eigenvalue

With the help of Eigenvalue we get to realize about the importance of the eigenvector. Each eigenvector is related to an Eigenvalue. Eigenvalues have really much importance in the PCA method. To keep just principal Eigen values, they will permit to filter the non-significant eigenvectors.

6) Eigenvectors

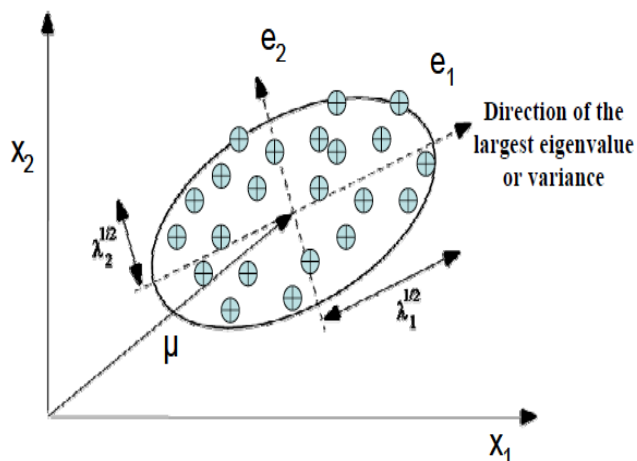
The eigenvector of a linear operator which we have to operate on the operator, gives a scalar multiple of themselves. Scalar is called the Eigenvalue associated with the eigenvectors

Orthogonal transformation is used in Principal component analysis (PCA) which is a statistical procedure. A observation set of correlated variables is converted into values forming a set of uncorrelated variables. They are called as principal components. PCA was introduced by:

1. Karhunen-Loeve (1947-1963)
2. Hotelling for psychometry (1933)
3. Pearson for biology (1901)

To remove correlation between variables or signals PCA is very useful analysis method, while it finds directions with maximal variance at the same time. Assume that we have access to N samples of a vector x with n elements. Pixel grey levels, values image features of a signal instants can be measurements of elements of x at different instants. These vectors will be scattered. In the n -dimensional space these vectors will not be uniformly distributed

If N data vectors of dimension n from the dataset are given, we have to find $c \leq N$ orthogonal vectors such that these vectors can be best used to represent data. Now we need to find an orthonormal basis that will maximize the variance of the projection of the dataset vectors along the new coordinate axis. In this first axis corresponds to the maximal variance; the second axis corresponds to the maximal variance in the direction orthogonal to the



1) Evaluation:

1. High dimensional data is converted into low dimensional image. In this principal components are not correlated, it is done in linear fashion.

2. When the number of test images increases the recognition rate of PCA perform better, but the rate of recognition reduces by certain number.
3. The number of probes images is larger to the number of reference images before PCA projection. Size of image is not superior issue.

VII. HU MOMENTS

With the help of geometrical moments of the hand region Hu invariants are calculated. The first six descriptors are translation invariant, scale invariant and rotation invariant. Skew invariance is ensured by seventh descriptor, which enables to distinguish between mirrored images. To provide a scale, orientation and position invariant representation of an object Hu invariants are used. In contrast to FD, moments are region-based descriptors.

VIII. CLASSIFICATION

For each class i , a training set is used to learn, corresponding to a gesture, a mean invariant vector m_i and a covariance matrix L_i . Then each image, represented by the invariant vector x , is classified by minimizing the distance. A Bayesian distance is used to classify hand shapes.

IX. CONCLUSION

A comparative study between various methods of hand gesture recognition is presented in this paper. The performance of various methods is tested and from the results we can conclude that particular method is superior to other. We have made acquisition of larger database, our own gesture vocabulary is defined. We could test the robustness of the PCA and Hu invariants to changes in scale, translation, rotation and point of view. Future work will aim at making improvement in the hand detection with the help of tracking methods, and taking benefit of the temporal stability of a given gesture.

REFERENCES

1. V. Pavlovic, R. Sharma, and T. Huang, "Visual interpretation of hand gestures for Human-Computer Interaction: A review," vol. 19, no. 7, pp. 677-692, Jul. 1997
2. F.-S. Chen, C.-M. Fu, and C.-L. Huang, "Hand gesture recognition using a real-time tracking method and Hidden Markov Models," vol. 21, no. 8, pp. 745-758, Aug. 2003
3. G. Heidemann, "The principal components of natural images revisited," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 5, pp. 822-826, May 2006