Hand Gesture Recognition Using PCA and Histogram Projection

Krishnakant C. Mule & Anilkumar N. Holambe

TPCT's COE ,Osmanabad, MH, India
E-mail : mulekc@gmail.com, anholambe@gmail.com

Abstract – **The recognition problem is solved through a matching process in which the segmented hand is compared with all the Images in the database.**

In this paper a novel approach for vision based hand gesture recognition is proposed by using both principal component analysis (PCA) and projection method for feature extraction. We conclude with future abilities.

Keywords – Hand gesture, PCA, Image projection.

I. INTRODUCTION

One of the main goals of Hand Gesture Recognition is to identify hand gestures and classify them as accurately as possible. For systems to be successfully implemented, it is critical that their performance is known. To date the performance of most algorithms has only been reported on identification tasks, which imply that characterization on identification tasks holds for verification [1][2].

In this paper will discuss an approach for man-machine interaction [3] using a video camera to interpret the American one-handed sign language alphabet and number gestures.

Sign languages: Sign languages are the natural communication media of hearing-impaired people all over the world [4]. Sign languages are well-structured languages with a phonology, morphology, syntax and grammar. They are different from spoken languages, but serving the same function. The aim in SLR is to reach a large-vocabulary recognition system which would ease the communication of the hearing impaired people with other people or with computers [2].

Hand gestures : Hand gestures are the independent way of communication. Hand gestures can be considered as complementary modality to speech. Gestures are consciously and unconsciously used in every aspect of human communication and they form the basis of sign languages.

II. HAND GESTURE RECOGNITION SYSTEM

The Hand gesture recognition process can be coarsely divided into four phases.

Flow of hand gesture recognition is shown below

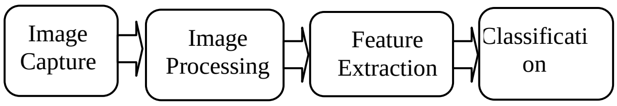

Fig1. Schematic view of gesture recognition process

A. Image Capture

The task of this phase is to acquire an image, or a sequence of images (video), which is then processed in the next phases.

B. Preprocessing

As Preprocessing prepares the image so as to extract the features in the next phase It is the process of dividing the input image (in this case hand gesture image) into regions separated by boundaries [5].The most commonly used technique to determine the regions of interest (hand), is 'skin color detection'[6].

C. Feature Extraction

Feature extraction is a form of dimensionality reduction. This finds and extracts features that can be used to classify the given gesture [7].

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (much data, but not much information) then the input data is transformed into a reduced representation set of features (also named features vector).

Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input [8]. Different gestures result in different, good discriminable features.

D. Classification

The classification represents the task of assigning a feature vector or a set of features to some predefined classes in order to recognize the hand gesture.

In general, a class is defined as a set of reference features that were obtained during the training phase of the system or by manual feature extraction, using a set of training images.

III. HGR USING HISTOGRAM PROJECTION AND PCA

In this paper will discuss how to use histogram projection as method of feature extraction for extracting feature pixels from the input image and then afterwards how to use Principal Component Analysis (PCA) [9] to reduce the size of feature vector.

A. Using Histogram Projection For Feature Extraction

We are using the histogram projections of the images to extract the features from the input image. The projection method includes the following steps.

i. Hand Detection

The $Y\ C_b C_r$ color space

The proposed hand region detection technique is applied in the $Y\ C_b C_r$ color space [5]. In particular, Y is the luminance component and C_b, C_r are the chrominance components [5]. RGB values can be transformed to $Y\ C_b C_r$ color space using the following equation [5]:

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Eq.1 RGB to $Y\ C_b C_r$ conversion

The classification of the pixels of the input image into skin color and non-skin color clusters is accomplished by using a thresholding technique [10] that exploits the information of a skin color distribution map in the $Y\ C_b C_r$ color space.

In this method, a map of the chrominance components of skin color was created by using a training set of images.

ii. Histogram Projection

Projection is one of the simple scalar descriptor[12]. Region description by projections is usually connected to binary image processing. Projections can serve as a basis for definition of related region descriptors; for example, the width (height) of a region with no holes is defined as the maximum value of the horizontal (vertical) projection of a binary image of the region. Projections can be defined in any direction.

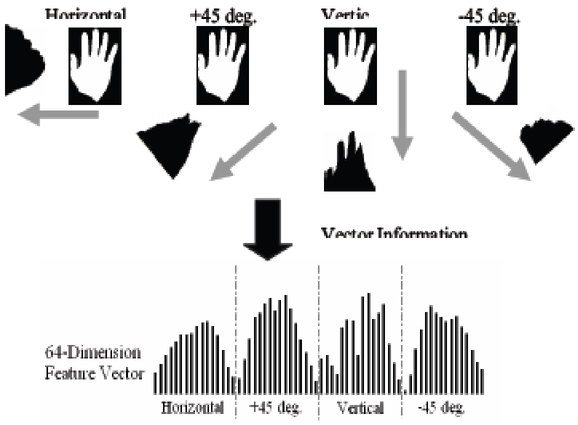

Fig. 2 : Hand feature vector extraction using different projections.

The first step is to count the pixels in four directions, i.e. horizontal, vertical, +45deg and -45deg directions from the handled image. Namely, horizontal, vertical, +45deg and -45deg directions pixels are projected in respective directions and reduced to feature vectors.

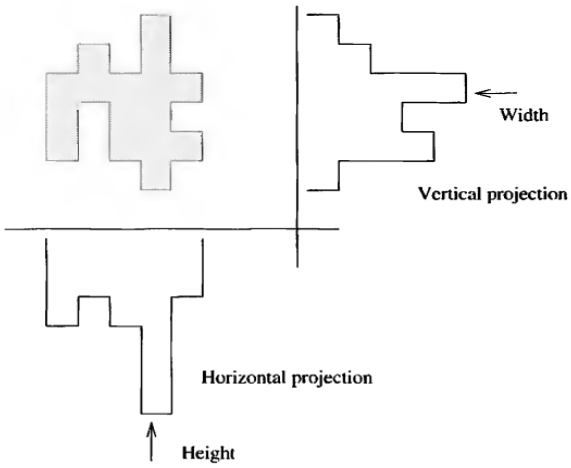

Fig.3: Horizontal and Vertical Projection

Finally, a vector is formed by concatenating the four vectors in the order of horizontal, +45 deg, vertical and -45 deg directions .The information by means of template vectors, calculates the similarity (= Euclidian distance: ED) between an input vector (hand gesture feature) and template vectors [11] (large testing/training data), and returns the maximal likelihood vector, and classify the different hand gestures [12].

B. Principal Component Analysis (PCA)

We can use PCA to compute and study the Eigenvectors of the different pictures and then to express each image with its principal components [9] (Eigenvectors).

It is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. First of all, we had to create the data set. The aim is to choose a good number of pictures and a good resolution of these in order to have the best recognition with the smallest database. Then, the next step is to subtract the mean from each of the data dimensions. The mean subtracted is simply the average across each dimension. The step three is to calculate the covariance matrix of the database. We could not calculate the covariance matrix of the first matrix, because it was too huge. So we had to find a way to find out the principal eigenvectors without calculating the big covariance matrix. The method consists in choosing a new covariance matrix. Our covariance matrix for A was called C and C is defined by [13]:

$$C = A^* A'$$

The Eigenvectors and the Eigenvalues of C are the principal components of our data set.

The PCA Algorithm :

Training phase:

Step 1: Obtain the training images $I_1, I_2, ..., I_M$

Step 2: Represent every image I_i as a vector G_i

Step 3: Compute the average image vector Ψ:

$$\psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i$$

Step 4: Subtract the mean image:

$$\Phi_i = \Gamma_i - \Psi$$

Step 5: Compute the covariance matrix C :

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^{\,T} = AA^T \qquad (\mathrm{N}^2 \times \mathrm{N}^2 \text{ matrix})$$

$$where, A = [\Phi_1\ \Phi_2\ ...\ \Phi_M] \quad (\mathrm{N}^2 \times \mathrm{M} \text{ matrix})$$

Step 6: Compute the eigenvectors u_i of AA^T

The matrix AA^T is very large. So, compute eigenvectors v_i of A^TA, which has same eigen values and eigenvectors.

Step 7: Compute the M best eigenvectors of AA^T : $u_i = Av_i$

Step 8: Keep only K eigenvectors (corresponding to the K largest eigenvalues)

Detection Phase:

Given an unknown image G,

Step 1: Compute: $\Phi = \Gamma - \Psi$

Step 2: Compute: $\hat{\Phi} = \Gamma - \Psi$

Step 3: Compute: $e_d = \left\| \Phi - \hat{\Phi} \right\|$

IV. EXPERIMENTAL RESULTS

The experimental results of hand gesture recognition by using the projection method and PCA method separately are given in Table 1.

For experimental purpose we used 10 hand gestures of 10 digits from 1 to 10 in ASL (American Sign Language) with 10 different variations. So we have database of totally 100 images for the experiment purpose.

In Table 2 we have the result of proposed method by combining the projection method and the PCA method.

For some gestures the combination of projection and PCA yields more accuracy than using the projection and PCA methods separately using MATLAB.

Gesture	Method	Total	Correct	Accuracy
1	Projection	10	9	90%
	PCA	10	8	80%
2	Projection	10	6	60%
	PCA	10	7	70%
3	Projection	10	10	100%
	PCA	10	10	100%
4	Projection	10	6	60%
	PCA	10	5	50%
5	Projection	10	9	90%
	PCA	10	8	80%
6	Projection	10	7	70%
	PCA	10	8	80%

7	Projection	10	10	100%
	PCA	10	8	80%
8	Projection	10	7	70%
	PCA	10	8	80%
9	Projection	10	9	90%
	PCA	10	9	90%
10	Projection	10	8	80%
	PCA	10	7	70%
Average	Projection	100	81	81%
	PCA	100	78	78%

Table 1: Individual results of projection and PCA methods

Gesture	Method	Total	Correct	Accuracy
1	Projection + PCA	10	8	80%
2	Projection + PCA	10	7	70%
3	Projection + PCA	10	10	100%
4	Projection + PCA	10	6	60%
5	Projection + PCA	10	8	80%
6	Projection + PCA	10	8	80%
7	Projection + PCA	10	9	90%
8	Projection + PCA	10	8	80%
9	Projection + PCA	10	9	90%
10	Projection + PCA	10	8	80%
Average	Projection + PCA	100	81	81%

Table 2: Experimental results combining projection and PCA methods.

V. FUTURE SCOPE

In this paper we have only considered the static gesture, but in real time we need to extract the gesture form the video or moving scene. Therefore the system needs to be upgraded to support dynamic gesture. Proposed system can be further upgraded to give order and control robots. It can also be very helpful for the physically impaired persons. Above method can be further enhanced for binary and color images.

Some more applications are that this proposed system can be used for gaming. Instead of using the mouse or keyboard, we can use some pre-defined hand gesture to play any game. Also, this system can be used to operate any electronic devices by just keeping a sensor which recognizes the hand gestures. Another application is that this can be used for security and authorization by keeping any particular hand gesture as the password.

VI. REFERENCES

[1] Joseph J. LaViola Jr., (1999). “A Survey of Hand Posture and Gesture Recognition Techniques and Technology”, Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA.

[2] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). “A Rotation Invariant Approach on Static-Gesture Recognition Using Boundary Histograms And Neural Networks,” IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura.

[3] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, Mo Nours Arab, (2008) .“Human-Computer Interaction: Overview on State of the Art”, International Journal on Smart Sensing and Intelligent Systems, Vol. 1(1)

[4] S. Mitra, and T. Acharya. (2007). “Gesture Recognition: A Survey” IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311-324, doi:10.1109/TSMCC.2007.893280

[5] N. Ibraheem, M. Hasan, R. Khan, P. Mishra, (2012). “comparative study of skin color based segmentation techniques”, Aligarh Muslim University, A.M.U., Aligarh, India.

[6] HAND GESTURE RECOGNITION VIA A NEW SELF-ORGANIZED NEURAL NETWORK,E. Stergiopoulou, N. Papamarkos* and A. Atsalakis

[7] Xingyan Li. (2003). “Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm”,Department of Computer Science. The University of Tennessee Knoxville.

[8] M. M. Hasan, P. K. Mishra, (2011). “HSV Brightness Factor Matching for Gesture Recognition System”, International Journal of Image Processing (IJIP), Vol. 4(5).

[9] A tutorial on Principal Components Analysis, Lindsay I Smith

[10] Mokhar M. Hasan, Pramod K. Mishra, (2012). “Robust Gesture Recognition Using Gaussian Distribution for Features Fitting’, International Journal of Machine Learning and Computing, Vol.2(3).

[11] Mokhar M. Hasan, Pramod K. Mishra, (2012) “Features Fitting using Multivariate Gaussian Distribution for Hand Gesture Recognition”, International Journal of Computer Science & Emerging Technologies IJCSET, Vol. 3(2).

[12] W. T. Freeman and Michal R., (1995) “Orientation Histograms for Hand Gesture Recognition”,IEEE International Workshop on Automatic Face and Gesture Recognition.

[13] Hand Gesture Recognition:A Comparative Study Prateem Chakraborty, Prashant Sarawgi, Ankit Mehrotra, Gaurav Agarwal, Ratika Pradhan

[14] Ibraheem, M. Hasan, R. Khan, P. Mishra, (2012). “comparative study of skin color based segmentation techniques”, Aligarh Muslim University, A.M.U., Aligarh, India.

[15] Malima, A., Ozgur, E., Cetin, M. (2006). “A Fast Algorithm for Vision-Based Hand Gesture Recognition For Robot Control”, IEEE 14th conference on Signal Processing and Communications Applications, pp. 1-4. doi: 10.1109/SIU.2006.1659822

[16] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). “Survey on Gesture Recognition for Hand Image Postures”, International Journal of Computer And Information Science, Vol. 5(3), Doi:10.5539/cis.v5n3p110

◈◈◈