GESTURE RECOGNITION SYSTEM

Satish Kumar Kotha[1], Jahnavi Pinjala[2], Kavya Kasoju[3], Manvitha Pothineni[4]

[1]*Department of Biomedical Engineering, GRIET, Hyderabad*
[2]*Department of Biomedical Engineering, GRIET, Hyderabad*
[3]*Department of Biomedical Engineering, GRIET, Hyderabad*
[4]*Department of Biomedical Engineering, GRIET, Hyderabad*

Abstract

This paper presents a novel approach for the gesture recognition system using software. In this paper the real time image is taken and is compared with a training set of images and displays a matched training image. In this approach we have used skin detection techniques for detecting the skin threshold regions, Principle Component Analysis (PCA) algorithm and Linear Discriminant Analysis (LDA) for data compressing and analyzing and K-Nearest Neighbor (KNN), Support Vector Machine (SVM) classification for matching the appropriate training image to the real-time image. The software used is MATLAB. The hand gestures used are taken from the American Sign Language.

Keywords— *PCA algorithm, LDA algorithm, skin detection, KNN and SVM classification*

----------***----------

1. INTRODUCTION

Gesture recognition is an area of current research in computer vision. Body language is one of the important ways of communication among the humans. Thus, gesture recognition system would be an ideal approach for improving human-machine interaction. This kind of human-machine interfaces will allow a human to control a wide variety of devices remotely through hand gestures.

In the earlier stages of research many have proposed different approaches for the sign or gesture recognition system.

Eng-Jon Ong and Bowden presented an approach not only detecting the presence of human hands within an image but classifying the hand shape too [1] by using grey scale image. The Korean Manual Alphabet (KMA) by Chau-Su Lee et al presented a vision-based recognition system of Korean manual alphabet [1] which is a subset of Korean Sign Language.

Danieal Collobert proposed the same using [2] Input-output Hidden Markov Models. Fujimura proposed hand gesture recognition [2] using depth data.

Tin Hninn Hninn Maung used a simple and fast algorithm using orientation histograms will be developed. Pattern recognition system will be using a transform that converts an image into feature vector [3], which then will be compared with the feature vectors of a training set of gestures.

Ankit Chaudhary, J. L. Raheja, Karen Das, Sonia Raheja used fuzzy logic, artificial neural network, genetic algorithms etc. The methods in the pre processing of image for segmentation and hand [4] image construction also taken into study. Klimis Symeonidis proposed a pattern recognition system , by using a transform that converts an image into a feature vector, and will be compared with the feature vectors of a training set of gestures [5]. The system will be implemented with a Perceptron network.

2. SYSTEM ANALYSIS

The real time image is taken through the web cam and the training sets of images are taken from the Marcel database. Then pre-processing of the real time image and the training set of image is done by skin detection. Then the PCA and LDA algorithm is applied to the training sets of images for compressing and analyzing the images, then the KNN and SVM classifications are used to classify the real time image with the correct match of the training set of image. The fig 1 shows the block diagram of the proposed Gesture Recognition System.

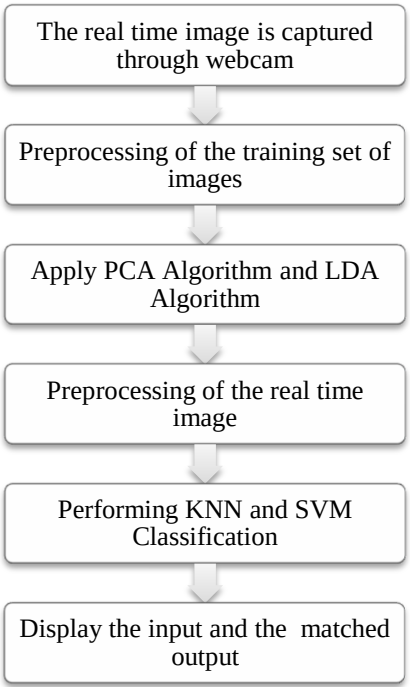


Fig 1: System Architecture

The real time image is taken through the web cam and winvideo supporting file should be downloaded from the math works for interfacing the webcam with the MATLAB. Image acquisition tool box of the MATLAB is used. The training sets of images are taken from the Marcel database and the size of the images are about 66 X 76 pixels.

Pre-processing: Skin detection is done as following

- After gathering the set of images, mark each pixel as "skin" or "non skin".
- Create an RGB histogram for "skin pixels" and another one for "non-skin pixels" (they are 32x32x32 in size).
- For a particular bin (pixel colour), the skin is log (H(r,g,b) / h(r,g,b)) ...where H stands for the skin histogram, and h is non-skin histogram.
- For a new image, compute the log likelihood of each pixel and then threshold the result to decide skin/non-skin.

PCA Algorithm: Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values [6] [7] of linearly uncorrelated variables called principal components.

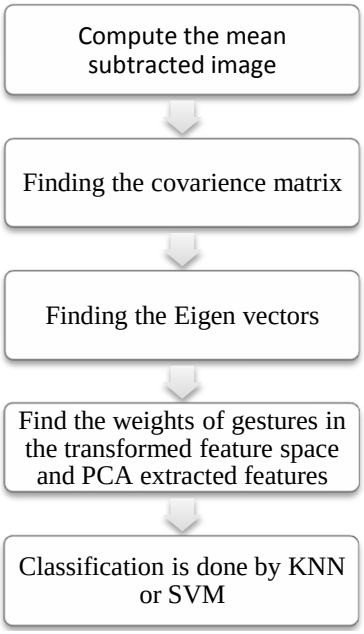


Fig 2: Block diagram of PCA algorithm

LDA Algorithm: Linear Discriminant Analysis (LDA) and related Fisher's linear discriminant are the methods used in pattern recognition [8], statistics and machine learning for finding a linear combination of features which characterizes [9]or separates two or more classes of objects or events. Then the resulting combination can be used more commonly, for dimensionality reduction or a linear classifier.

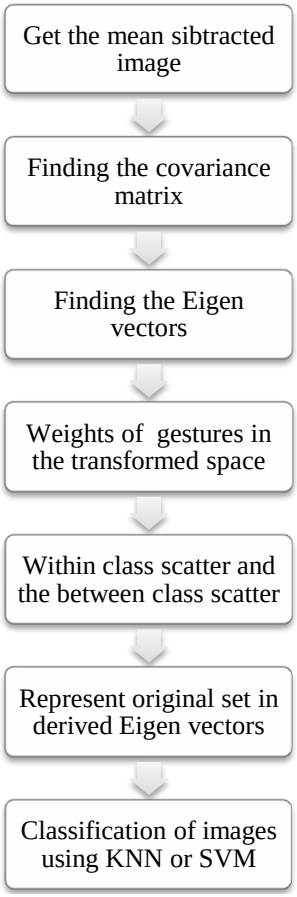


Fig 3: Block diagram for LDA algorithm

KNN Classification: The k-nearest neighbor algorithm is a classifying method which classifies an object where the majority of the neighbor belongs to. The choice of the number of neighbors is discretionary and up to the choice of the users. If k is 1 then it is classified [10] whichever class of neighbor is nearest.

Typically the object is classified based on the labels of its *k* nearest neighbors by majority vote. If *k*=1, the object is classified as the class of the object nearest to it. When only two classes are present, it is said that *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, we used the most common distance function for KNN which is Euclidean distance.

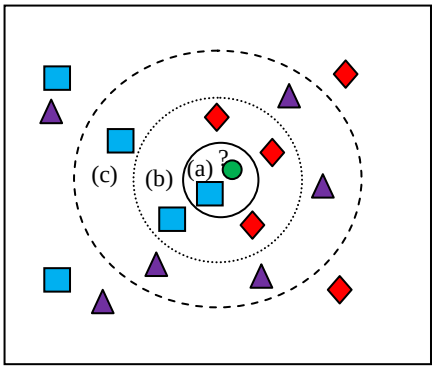


Fig 4: KNN classification. At the query point of the circle depending on the k value of 1, 5, or 10, the query point can be a rectangle at (a), a diamond at (b), and a triangle at (c).

SVM Classification: A support vector machine (SVM) is a non probabilistic linear binary classifier, which can analyze input data and predict which of the two classes it belong to.

It works by building a hyperplane separating the two classes which is of higher dimension.

A good separation is obtained by a hyperplane that is very far from any data point of each class [11], since further the separation of the data, better the performance.

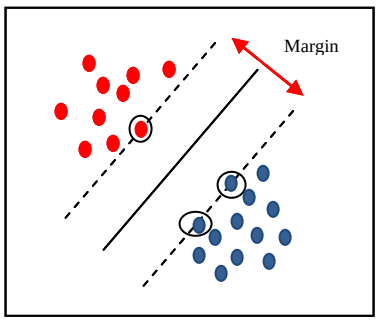


Fig 5: SVM Classification. In multidimensional space, support vector machines find the hyperplane that maximizes the margin between two different classes. Here the support vectors are the dots circled.

3. PERFORMANCE ANALYSIS

The performances of the proposed system according to the PCA and LDA analyses has mentioned below.

Fig 6 shows the alphabets used in this paper.

Fig 6: Alphabets used

These images are used to extract the features using performing some processing like skin extraction (fig 7).

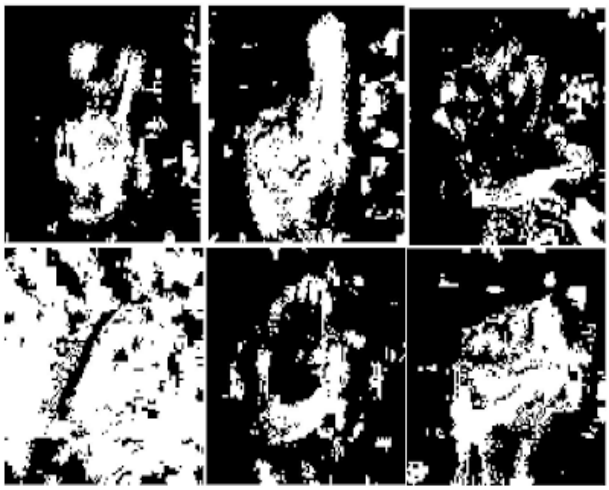


Fig 7: Results of skin thresholding

The extracted Eigen vectors are shown in the fig 8.

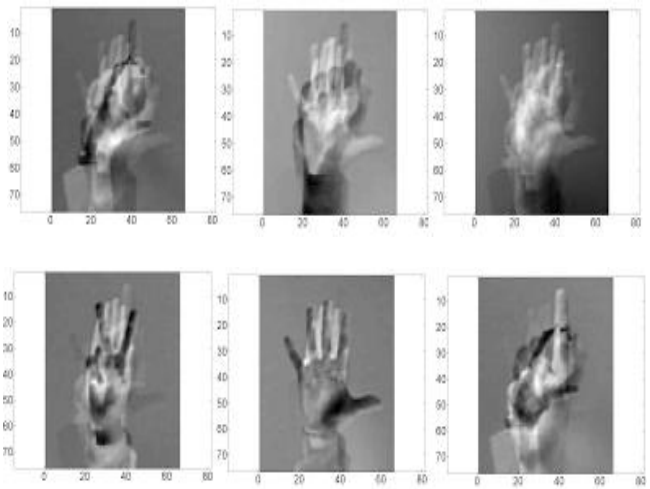


Fig 8: Eigen vectors

Using these Eigen vector defined subspace, we represent all the images in this subspace. Based on these representations we classify the Testing database images using KNN and SVM classifiers.

PCA Algorithm:

Fig 9: Classification using KNN

Fig 10: Classification using SVM

The results were observed in the following table 1:

Table 1: Accuracy of PCA

Accuracy %	KNN	SVM
Training Set	75.54	55.93

LDA Algorithm:

Fig 11: Classification using KNN in LDA

Fig 12: Classification using SVM in LDA

The results for this are given in the table 2:

Table 2: Accuracy of LDA

Accuracy %	KNN	SVM
Training Set	24.5	24.3

4. RESULTS

The tabular column corresponding to the results obtained using the SVM (table 4) and KNN (table 3) classifications. The input image is taken from the web camera and in this project we have used only 6 gestures for testing.

Table 3: Results of KNN classification

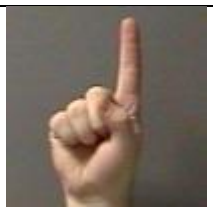
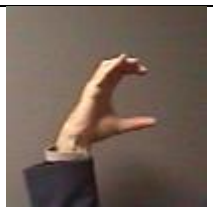
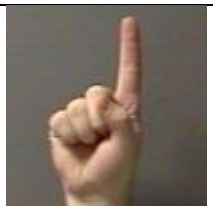
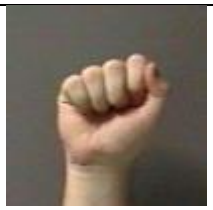
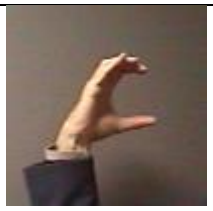
GESTURE	Without imcrop		Complete white Background	
	PCA	FLD	PCA	FLD
 5	5	5	5	5
 V	5	5	V	5
 Point	Point	5	Point	5
 A	Point	5	Point	5
 B	5	5	5	5
 C	5	5	C	5

Table 4: Results of SVM classification

GESTURE	Without imcrop		Complete white Background	
	PCA	FLD	PCA	FLD
 5	B	5	5	5
 V	Point	5	V	5
 Point	Point	5	Point	5
 A	Point	5	V	5
 B	V	5	5	5
 C	C	5	C	5

5. APPLICATIONS

Hand gesture recognition is been applied in different domains with different applications.

1. Sign Language Recognition: For the deaf and dumb people to communicate [12] through the sign language.
2. Robot Control: Controlling the robot using gestures for [12] example, “one” means “move forward”, “five” means to “stop”, and so on.
3. Graphic Editor Control: Graphic editor control system requires the hand gesture to be tracked and located as a pre-processing operation used 12 dynamic gestures [12] for drawing and editing graphic system.
4. Virtual Environments: For communication media system for 3D pointing [12] gesture recognition.
5. Numbers Recognition: Recognising meaningful gesture [12] using the Arabian numbers from 0 -9.
6. Television Control: Controlling the volume, changing the channels etc [12] can be done for using the gesture recognition.
7. 3D Modelling: Building the 3D models by showing the models [12] through the hand gestures.

6. CONCLUSION

This project developed a system that can recognize real time image based on the features we extracted from the training database using Principal Component Analysis and Linear Discriminant Analysis algorithms. These were classified using Support Vector Machine, K Nearest Neighbor, and Neural Network schemes.

There are few reasons for poor performance of testing data. Firstly, the system faced problems in the skin detection. Because there is no strict threshold set for skin probabilities, some of the background is also in skin color, so it may be mistakenly detected as skin. The variation in images also played a major part in the low accuracy, which includes different background of images, different sizes of gestures in the images, different orientations and angle of gestures, etc. The different illumination effects for different images too posed a problem.

FUTURE SCOPE

We further would like to extend it by giving the hand gestures meaning in the form of text or by voice.

For extension we would like to use the Artificial Neural Networks and the Image processing tool box in the MATLAB and other sources for the conversion of the gesture meaning into the text format or by the voice.

REFERENCES

[1] O. M. Foong, T. J. Low and S. Wibowo, "Hand Gesture Recognition: Sign to Voice System (S2V)," 2008

[2] P. Chakraborty, P. Sarawgi, A. Mehro, G. Agarwal and R. Pradhan, "Hand Gesture Recognition: A Comparative Study," in International MultiConference of Engineers and Computer Scientists, Hong Kong, 2008.

[3] T. H. H. Maung, "Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks," 2009.

[4] A. Chaudhary, J. L. Raheja, K. Das and Sonia, "Intelligent Approaches to interact with Machines using Hand Gesture Recognition in Natural way: A Survey," International Journal of Computer Science & Engineering Survey (IJCSES), vol. 2, 2011.

[5] K. Symeonidis, "Hand Gesture Recognition Using Neural Networks"

[6] K. Kosmelj, J. Le-Rademacher and Lynne, "Symbolic Covariance Matrix for Interval-valued Variables and its Application to Principal Component Analysis: a Case Study," in Metodoloski zvezki, 2014.

[7] "A tutorial on PCA," [Online]. Available: http://www.sccg.sk/~haladova/principal_components.pdf.

[8] "Fisher Linear Discriminant Analysis," [Online]. Available: http://www.ics.uci.edu/~welling/classnotes/papers_class/Fisher-LDA.pdf.

[9] Wikipedia, "Linear discriminant analysis," [Online]. Available: http://en.wikipedia.org/w/index.php?title=Linear_discriminant_analysis&oldid=153494000.

[10] J. Kim, B.-s. Kim and S. savarese, "Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines".

[11] Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/Support_vector_machine.

[12] P. K. K. Vyas, A. Pareek and D. S. Tiwari, "Gesture Recognition and Control Part 2 – Hand Gesture Recognition (HGR) System & Latest Upcoming Techniques," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 1, no. 8.

[13] R. Sharma, Y. Nemani, S. Kumar, L. Kane and P. Khanna, "Recognition of Single Handed Sign Language Gestures using Contour Tracing Descriptor," in Proceedings of the World Congress on Engineering, 2013.

[14] Prof.Sayyad T.J, "Recognition of ASL using Hand Gestures," International Journal of Emerging Trends & Technology in Computer Science (IJETTCS, vol. 2, no. 5, 2013.

[15] Malima, E. Ozgur and M. Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control," IEEE, 2006.

[16] X. Li, "Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm"

[17] S. K. Kang, M. Y. Nam and P. K. Rhee, "Color Based Hand and Finger Detection Technology for User Interaction," 2008

[18] R. J, K. Mahesh, S. Mahishi, D. R, S. S and N. V. Pujari, "“Finger Detection for Sign Language Recognition," in Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, 2009

[19] S. D.Gawande and P. N. .. Chopde, "Neural Network based Hand Gesture Recognition," International Journal of Emerging Research in Management &Technology, vol. 2, no. 3.

[20] "Automatic sign language detection," [Online]. Available: https://sites.google.com/site/autosignlan/.