Research Article
Hand Posture Recognition Human Computer Interface

[1]Abida Sharif, [1]Saira Latif, [2]Muhammad Irfan Sharif and [2]Mudassar Raza
[1]Department of Computer Sciences, COMSATS Institute of Information
Technology Islamabad, 44000, Pakistan
[2]Department of Computer Sciences, COMSATS Institute of Information
Technology Wah Cantt., 47040, Pakistan

Abstract: The basic motivation behind this research work is to assist enormous number of disable people to enhance their capabilities regardless of their disability. In particular, we have focused and addressed the deafness and dumbness disability in human beings on technological basis. The aim was to design a system to help such people. Sign language has been used as our database in whole recognition process. The gestures have been read by comparison with available signs in the database. The work is comprised of three major parts. 1) Acquiring images in real time environment through any imaging device. 2) Recognition of those images on the basis of probability by comparing with the database. 3) Finally translating recognized images into possible output. We have used various algorithms to validate the approach and to check the efficiency. In particular, mainly adaboost and Support Vector Machine (SVM) algorithms have been tested. Both of these algorithms worked well but SVM was found to be optimum with respect to time efficiency as compared with adaboost.

Keywords: Adaboost, database, hand posture recognition, skin detection, SVM

INTRODUCTION

Biometrics is emerging as a leading technology now days. The most important thing in biometrics is to recognize the features (Sharif *et al*., 2011) like, Face (Sharif *et al*., 2012), Finger Prints, Iris and Gait etc. Technological revolution in science and engineering has helped human beings to overcome their weakness and make their existence more valuable. Everyone is born with some abilities along with some disabilities too. An aspect of all human thinking up till now was to use the abilities for the development of human society in order to make the people's life more useful irrespective of their disabilities. Communication is the basic attribute of human beings. Normally everyone communicates in its natural way using its natural senses i.e., speaking with tongue and listening with the ears together with body language and gestures. Consequently, persons born deaf and dumb encounter very serious problems. Gestures were traditionally used by such people as an alternative mean of communication. However in the modern era scientists have tried to realize better solutions to cope with this issue. The present work is another such effort in the series of such attempts. We conceived and developed a system to recognize the hand gestures and read them into words or actions. The proposed work can be divided into three parts. First, the system takes the real

time hand images and extracts the gesture. Secondly it recognizes the posture and finally converts into possible action. Different classification algorithms have been successfully tested experimentally for the sake of efficiency and robustness of the whole scheme. The experimental results verified the validity of the model. The realized system can have a clear recognition in real time with minimum error rate.

In paper (Pansare *et al*., 2012) authors intended a real time hand sign scheme. In this scheme experimental setup based on low-cost web camera used for taking snapshot only using Red Green Blue [RGB] color. This structure consists of four stages at first stage detained RGB image converted into binary image using gray threshold technique. At second stage crop the hand area and then Sobel filter is used for edge detection of the cropped hand. Third stage creates feature vector as centroid and area of edge. At fourth stage this feature vector compared with feature vector a training dataset of gestures. Faultless matching presents ASL alphabet, significant word using file handling.

In paper (Rautaray and Agrawal, 2012), proposed a hand gesture recognition system which is worked for images browsing in the image browser and make available productive results in the direction of user friendly interface involving human and computer by means of hand gestures. This proposed system based on three major parts hand segmentation, hand tracing and

---

gesture recognition from hand feature. The proposed scheme further incorporated with unlike applications similar to image browser, virtual game etc., potential for human computer interaction.

Paper (Hassanpour and Shahbahrami, 2009) is the survey in which authors discusses the different techniques of analyzing, molding and recognizing hand postures in the perspective of the Human-Computer Interface (HCI). Classification of the different algorithms consists on the uses that they have been build up for and comes in the reach that they have worked to differentiate the presented postures. As well give direction for future development.

Ghotkar and Kharate (2012) introduced visualization based hand posture identification system for HCI. Hand tracking and segmentation are the most important parts for hand posture identification system. This proposed system built a robust and well-organized hand tracking and segmentation Algorithm. This algorithm is more competent to tackle the different challenges related to vision based system like skin color exposure, complex background elimination and variable lighting situation.

Messer (2009) introduced static hand posture identification, which mostly based on the identification of well defined signs consist on a pattern of the hand.

Rokade *et al*. (2009) introduced RGB segmentation system which is supplementary responsive to light situations and the threshold value for adaptation of output image to binary image that value is unusual for different lighting environment.

MATERIALS AND METHODS

We have worked on two types of database of sign language. One is grey scale database contains grey scale images with black background and the second is automated database contains colored images with black background Limitations: A common constrain on both types of databases is that hand should not be connected with other parts of the body.

Grey scale database: In grey scale database we have 400 images. Some grey scale images shown in Fig. 1.

Filters use for hand extraction: For hand extraction we have three types of filters (Haider *et al*., 2012):
- Sobel filter
- Prewitt filter
- Canny Filter

Sobel filter: Two 3×3 convolution mask are applied to each and every pixel of the image in three ways, one with color, second in Horizontal direction and thirdly in vertical direction at a time. The outcome of each one convolution is indulgence a vector, which signifies the edge through the existing pixel. If the magnitude of the total of these two orthogonal vectors is beyond from user-defined threshold, the pixel is obvious in black as an edge. Otherwise, the pixel is set to white.
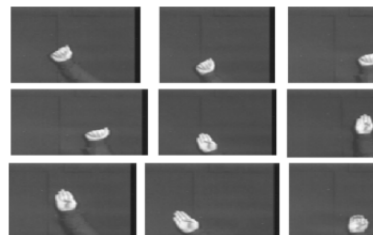

Fig. 1: Automated or grey scale database

Fig. 2: Prewitt filter, sobel filter and canny filter

Fig. 3: Image from automated database

Prewitt filter: Prewitt filter generates image where it found sharp changes in grey level values (edges) in image.

Canny filter: The canny filter makes use of multi-stage algorithms to find the edges in image in extensive range.

We check all the filters like Prewitt filter, Sobel filter and canny filter for extracting hand edges from the image of database. Results of Prewitt and Sobel filter is almost same but the canny filter gives too much edges that's why we cannot use canny filter because we need only hand edges. The results of Prewitt and Sobel filters are almost same, we can use any one. In our project we use Sobel filter for edge extraction from the images. Figure 2 shows the result of these three filters.

Steps of hand extraction: The next step is separate a hand from other areas like face arms etc. For this we use following techniques:

- First we take an original image from database (Fig. 3).
- Then we extract the edges of original image using Sobel filter (Fig. 4).
- Brighten the above extracted images to connect the edges by convolution (Fig. 5).
- We find the area of the connected regions in the image. The area of the line will be the maximum and the 2nd maximum area will be offhand by

Fig. 4: Edges of the original image using sobel filter
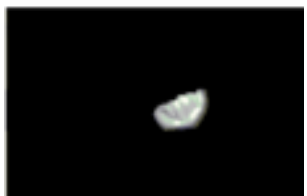

Fig. 5: Edge connection

Fig. 6: Extracted hand

Fig. 7: Extract filled hand

Fig. 8: Real database

intuitively. So we will put a check and will take 2nd maximum connected area. The second maximum area of the connected region will be offhand. Figure 6 represents the resultant extracted hand image.

- After multiplying the extracted hand with original image we extract filled hand (Fig. 7).

Real database: We take 200 hand images of 20 people including both male and female, in which each person has 10 images. These 20 people have different skin color and different hand structure but with black background. Figure 8 represents some images from the Real Database.

Hand extraction in real database: We extracted hand by using cropping technique. Cropping was done by two methods:

- Manual cropping
- Auto skin detection using colors thresh-holding technique.

Manual cropping: We initially cropped images by using built in command in MATLAB. We cropped images using this but we encounter with the problem of large size of image due to including surrounding area of hand which results in consuming more time during making model. To cope with this problem, we use auto skin detection method.

Auto skin detection: We generate a Code of auto skin detection and auto cropping. We simply input the real image taken through camera to the function of Auto Skin Detection code and it returned us hand excluding everything else in this purposed work two thresholding techniques are checked for the detection of skin. These are YCbCr with hue thresholding and RGB thresholding. The distinction among YCbCr with Hue thresholding and RGB thresholding results is that YCbCr shows color the same as brightness distinct signals of two colors and shows color the same as brightness, while RGB shows color as red, green and blue. In YCbCr, the Y indicates the brightness (luma), Cb indicates blue minus luma (B-Y) and Cr represents red minus luma (R-Y). YCbCr permits image compression techniques to get benefit that the eye is more judicious of brightness intensity than the color. Therefore, YCbCr liable to mostly supported for storing figures mainly photograph pictures and video. The YCbCr color model can supports TIFF and JPEG file format.

RGB thresholding: The RGB color representation is a preservative color representation in which red, green, blue light are mixed together in different manners to make an extensive array of colors. The name of this representation comes up to the initials of the three preservative main colors as red, green and blue:

- Thresholding value for fair color detection:

$$\{((R > 95)\&\&(G > 40)\&\&(B > 20)\&\&$$
$$(max(R, max(G, B)) - min(R, min(G, B)$$
$$)) > 15\&\&abs(R - B) > 15\&\&abs(R - G)$$
$$> 15\&\&R > G\&\&R > B\&\&abs(R - G)$$
$$< 100\&\&abs(B - G) < 105) \mid\mid (R > 220\&\&$$
$$G > 210\&\&B > 170\&\&abs(R - B)$$

(a)　　　　　(b)　　　　　(c)

Fig. 9: Skin detection using RGB thresholding

(a)　　　　　(b)　　　　　(c)
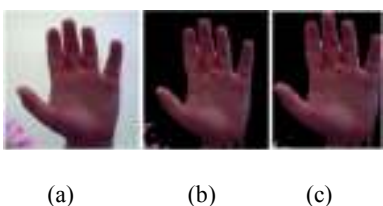
Fig. 10: Skin detection using Ycbcr and hue thresholding

$$> 15 \&\& abs(G - B) > 30 \&\& abs(R - G) \\ <= 15 \&\& R > B \&\& G > B)\} \qquad (1)$$

- Thresholding value for dark color detection:

$$((0.3 \degree R) > 0 \&\& (0.13 * R - 0.105353 * G > \\ 0 \&\& (0.108861 * R - 0.2 * B) > 0 \\ \&\& abs(R - G) < 70) \qquad (2)$$

- Thresholding value for union fair and dark color:

$$(((R1 \&\& G1 \&\& B1) \| (R2 \&\& G2 \&\& B2)) \sim= 0)$$

Results for RGB thresholding: Figure 9a shows the original image and Fig. 9b shows the image on which RGB thresholding applied. And the result is not clear, skinny area is clearly detected. The area where the red color is in greater quantity not be detected properly. Figure 9c shows the cropped image of hand.

YCbCr and hue saturation thresholding: YCbCr represent colors as a combination of following three values:

- Y shows the Brilliance (generally called the luminosity and brightness).
- Cb indicates the color of the blue mainly.
- Cr indicates the color of the red mainly (Green is attained by using a mishmash of these three standards).

Combine the thresholding of Hue, Cb, Cr for skin detection. The thresholding values are:

$$140 <= cr \& cr <= 165 \& 140 <= cb \& cb \\ <= 195 \& 0.01 <= hue \& hue <= 1 \qquad (4)$$

Cb, Cr and Hue more accurate to the result in real time. So we use combination of YCbCr and Hue.

Results for combination of hue, Cb and Cr for skin detection thresholding: Figure 10a shows the original image and Fig. 10b shows the image on which YCbCr and hue thresholding applied. And the result is clear, skinny area is clearly detected. Figure 10c shows the cropped image of hand.

LEARNING ALGORITHMS

We have selected following two algorithms for the classification of images because of their best results of classification among all other algorithm:

- Adaboost
- SVM

Adaboost: Adaboost, short for Adaptive Boosting, is a machine learning algorithm.

Adaboost uses weak learner. At the start it selects the learner that classifies more data properly. After that, the data is re-weighted to amplify the value of misclassified samples. This practice goes on furthermore; at every step the weight of all week learner along among other learners is decided. So, Adaboost classifier needs too much time for making training model. Adaboost cannot give accuracy with small training set so we moved to SVM.

SVM: SVM (Support Vector Machine) is an organized learning method applied for classification and regression. In other words, provided a set of training examples, each fitting in to one or the other kind, the SVM training algorithm creates a model that allocates the new example to first or second category.

In fact, an SVM training model is a demonstration of the examples that are mapped as spots in space; as a result the examples of the split categories are separated by a clear cut gap that is as broad as is achievable. New examples are then predicted and assigned to the group depending on which side of the gap they lie.

In simple words, SVM creates a hyper plan or set of hyper planes in high dimensional space, this hyper plane can be exploited for categorization, regression or in other tasks. By instinct, a fine division or separation is attained by that hyper plane that has major distance (function margin) to the closest training data of any class. In general, increases the distance decreases the generalization error of the classifier.

RESULTS AND DISCUSSION

We made a training set of 60 extracted hands from automated database. Then by using training algorithm we built a training model of these images. Then we tested an image from automated database by using

Table 1: Adaboost and SVM classifiers results

	Automated/grayscale database			Real database		
	Correct results (%)	Not predicted (%)	Wrong results (%)	Correct results (%)	Not predicted (%)	Wrong results (%)
Adaboost	80	20	20	58.33	25	41.66
SVM	95.5	5	5.66	90.40	5	10.66

testing algorithm along with this training model of adaboost. It gave us 80% accuracy with 16.66% error. But when we tested a real image of this training model, it showed error and less accuracy which was about 60%. To solve this problem we increased the size of training model. Then we tested again but the problem remained. Due to color difference in training and testing sets, we move toward formation of real database. Initially we made a database of 60 real images of 20 different people then we extracted hand from those images and made a training set and then perform testing which showed 60% accuracy with little more computational time. To solve this problem we increased the size of training set repeatedly 60 to 120, then 120 to 160, 160 to 190 up to 250, but the problems of less accuracy and more executions time still remained. Finally we move toward SVM. The same training set (which we prepared for adaboost) was trained on SVM and it gave us training model with in no time as compare to adaboost.

Then we tested a real image on this training model it gave us almost best results with the accuracy of 80%. Table 1 shows the experimental results of automated and real database hand posture recognition using Adaboost and SVM classifier.

CONCLUSION

In this study we discuss our purposed work which is divided into three parts. First, the system takes the real time hand images and extracts the gesture. Secondly, it recognizes the posture and finally converts into possible action. We check our results on Adaboost and SVM algorithms as shown in Table 1 by using training model of automated/grey scale images and real database. The results are not shown accurate. Adaboost algorithm is not efficient for small database. It has bad accuracy rate and more time consuming in making training model. While SVM algorithm is efficient than Adaboost. We have checked the Adaboost algorithm for 200 images on automated and real database. By proceeding above work on large scale, there are infinite number of possibilities of innovation and creativity .By

being successful in making large database we can get good results of adaboost too. We can use above work in tracking of objects. Some other considerations are needed in its error rate minimization and time efficiency.

REFERENCES

Ghotkar, A.S. and G.K. Kharate, 2012. Hand segmentation techniques to hand gesture recognition for natural human computer interaction. Int. J. Human Comp. Interac., 3(1): 15.

Haider, W., M.S. Malik, M. Raza, A. Wahab, I.A. Khan, U. Zia and H. Bashir, 2012. A hybrid method for edge continuity based on Pixel Neighbors Pattern Analysis (PNPA) for remote sensing satellite images. Int. J. Commun. Netw. Syst. Sci., 5(29): 624-630.

Hassanpour, R. and A. Shahbahrami, 2009. Human computer interaction using vision-based hand Gesture recognition. J. Comput. Eng., 1: 21-30.

Messer, T., 2009. Static Hand Gesture Recognition. University of Fribourg, Switzerland.

Pansare, J.R., S.H. Gawande and M. Ingle, 2012. Real-time static hand gesture recognition for American Sign Language (ASL) in complex background. J. Signal Inform. Process., 3: 364-367.

Rautaray, S.S. and A. Agrawal, 2012. Real time multiple hand gesture recognition system for human computer interaction. Int. J. Intell. Syst. Appl., 4(5): 56.

Rokade, R., D. Doye and M. Kokare, 2009. Hand gesture recognition by thinning method. Proceeding of the International Conference on Digital Image Processing, pp: 284-287.

Sharif, M., S. Mohsin and M.Y. Javed, 2011. Real time face detection using skin detection (Block Approach). J. Appl. Comp. Sci. Math. Suceava, 10(5).

Sharif, M., M.Y. Javed and S. Mohsin, 2012. Face recognition based on facial features. Res. J. Appl. Sci., 4(17): 2879-2886.