ISSN No. 0976-5697
Volume 2, No. 2, Mar-Apr 2011
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

Hand Gesture Recognition using skeleton of Hand and Hausdroff Distance

P.Swarna Latha *
School of Computing Science and Engineering,
VIT University, Vellore,
Tamil Nadu, India
pswarnalatha@vit.ac.in

M.Rajasekhara Babu
School of Computing Science and Engineering,
VIT University, Vellore,
Tamil Nadu, India
mrajasekharababu@vit.ac.in.

K.Siva Rajesh Reddy
School of Computing Science and Engineering,
VIT University, Vellore,
Tamil Nadu, India
ksivarajeshreddy2009 @vit.ac.in.

Abstract: This paper discusses the use of image processing and computer vision concepts in the interpretation of hand gestures. Gestures are non verbal communication .By using gestures we can convey useful information or instructions to the machine (computer) .This is known as Human machine interaction (Human computer interaction(HCI).Hand gestures are an ideal way of exchanging information between human and computer, robots, or any other device . In this paper we use the skeleton of the hand instead of the entire hand, because of its robust nature against translation, rotation and scaling. Skeleton is computed for each and every hand gesture and superimposed on a single image called as Dynamic Signature of the particular gesture type. Gesture is recognized by using the Hausdroff distance measure by comparing the current Dynamic Signature of the particular gesture with the gesture Alphabet set.

Keywords: Static, Dynamic Gesture; Skeleton; Local Orientation Histogram (LOH); Hausdroff distance.

I. INTRODUCTION

The main goal of this paper is to direct some devices in industries or homes by using our hand gestures from a particular distance instead of going near to that device. For example if we want to turn off the machine in a industry from a certain distance by using our hand gestures. This hand gesture will control the machine by sending the some information to the machine; this information will be conveyed by hand gesture. For identifying hand gesture we are making use of the skeleton concept and image Euclidean distance measure.

A .Hand Gesture

Human Machine Interaction (HMI) or Human Computer Interaction (HCI) having utmost importance in our daily lives. Gesture recognition can be termed as one of the approach in this direction. It is the process in which the gestures made by some user or sender are recognized by the end user like machine or human. Gesture recognition and classification is one of the interested areas in computer vision domain.

Gestures are meaningful physical body movements of fingers, hand, arm or face. Gestures are classified as 1.Facial gestures 2.hand Gestures 3.Combination of hand and face gestures. In this paper we mainly concerns on hand gestures.

From the initial intention to the final performance, gesture follows a motion in space and time [1].Kendon distinguishes the hand gesture into three phases. I.e. a single gesture is made of 1. Preparation 2.Stroke 3.Retraction.Among this three phase's stroke carries meaningful information than the other two phases. The other two phases mainly consist of moving the arms from the beginning to the rest position and vice versa.

Hand gestures could be classified as a Static Hand gestures and Dynamic Hand gestures. Static Hand gestures are characterized by the hand posture which is determined by a particular palm-finger configuration. While Dynamic Hand gestures Characterized by start and end stoke hand configuration and general stroke motion [1].

B .Analysis of hand gesture

In this paper analysis of hand gesture is done by vision based technique. Vision based analysis is a natural and difficult one because of limitations of machine vision. Some of the limitations are segmentation of moving hand against a complex background, tracking of the hand, Skin color of hand for recognition. In this paper properties of an image (intensity values) are used for recognition of hand.

C. Classification of gesture recognition

Gesture recognition can be divided into two tasks; 1.Feature Extraction 2.Classification. Feature extraction mainly uses the low level information of an image and generates high level information for classification. Features are calculated based on distance, velocity, area, centroid, energy information and angle information [2, 3, and 4]. In this paper angle feature was used for static gestures and Image Euclidean distance (IMED) distance measure for dynamic gestures.

In this paper Conventional template matching [5] was used for classification. In conventional template matching templates are created by taking the average of raw data values of same gesture type with different samples. Image Euclidean Distance (IMED) was used for classification of hand gesture.

II. OVERVIEW OF PROPOSED APPROACH

Skeleton concept was used for dynamic hand gesture recognition and local orientation histogram for static gesture recognition. In this paper we are considering some assumptions (Constraints) such as.

1. The distance between the camera and hand is fixed.
2. Gestures are performed against a black background or uniform background.
3. Gestures are performed in the known region of space.
4. Taking only the Hand images no other body parts for our simplification.
5. Gesture alphabets are priori known to the user.

Generally gesture is a sequence of images (posture or frame).Instead of taking all the images considering only some sample of images(called as dynamic gestures) that will cover the entire motion of gesture, due to this the unwanted information and processing time will be eliminated.. For identification of a particular gesture some start and stop images (posture or static hand gesture) are used. After identification of start posture then only the dynamic hand gesture starts and ends only after identification of stop images. For the identification of start and stop images Local Orientation Histograms (LOH) are used because of the simple and fast calculation. The computation time is less because of dealing directly with pixel values rather than any training techniques.

A. Preprocessing step

This will be done before carrying out the actual processing. In this we mainly concerns on removing of the noise. RGB normalization reduces the intensity variations and removes shadows present in the image. Adaptive segmentation of the hand image is carried out because the color of the human hand varies from person to person. Hole filling of hand part also require for filling of the small dots present in the hand area. Blob removal is necessary for deletion of the some small dots in the hand image. These all preprocessing steps are required because of Skeleton mainly depends on the Segmented Hand.

B. Feature Extraction

Here we are mainly using two techniques.

1. Local Orientation Histogram :

LOH mainly used for recognition of static gestures (start and stop gestures).It mainly works based on the algorithm proposed by freeman and Roth [6].

2. Dynamic Signature:

Dynamic signature means the superimposition of skeleton images on a single image. The Dynamic signature gives the entire motion on a single image. Skeleton is robust against translation rotation and scaling.

C. Training Stage

In order to perform the gesture classification the gesture notations are already known. A set of samples are required for each gesture.

D. Classification Step

The current dynamic signature feature values are compared with database feature values the best match will lead to recognition of the unknown gesture.

III. IMAGE PREPROCESSING

The accuracy of the gesture recognition mainly depends on the preprocessing steps. For extracting the hand part background subtraction was used. After histogram based segmentation applied on image for extracting the hand part. Hole filling algorithm used for removal of the black spots in the hand part. Then the Small white dots removed by using blobs removal algorithm based on the area of the objects. If the object size is less than a particular threshold value then that objects removed from the image.

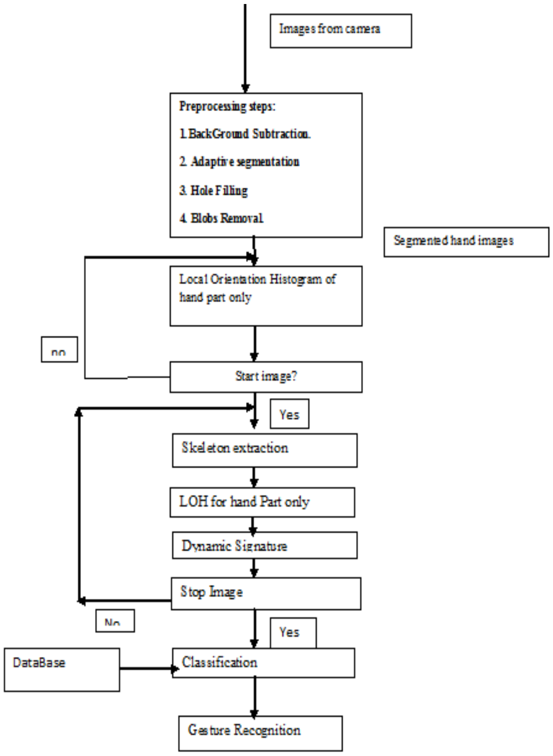
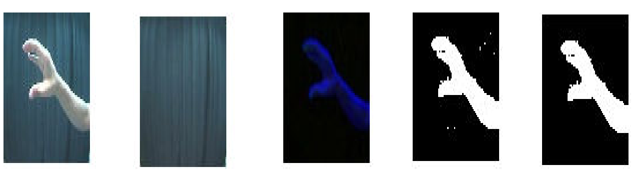


Figure 1: Gesture Recognition Diagram



a b c d e

Figure 2: a. Original Image. b. Background image. c. Background subtraction d. Threshold e. Blob Removal.

IV. STATIC GESTURE RECOGNITION

LOH used for static hand gesture recognition. LOH mainly used for recognition of start and stop gesture. The orientation histograms are comparatively easier and faster for computation.

A. *Local Orientation Histogram*

LOH [6] are also easier and faster for computation. LOH are robust against illumination changes. LOH used as features for each gesture and applied on the hand part of the gray part of the original image. The gray level image having hand part against a uniform black background. The local orientations histogram name itself indicates that directions of local gradients. The gradients are obtained by applying the sobel masks along the x and y directions. The sobel masks are

x-direction {0,1,2,-1,0,1,-2,-1,0}
y-direction {-2,-1,0,-1,0,1,0,1,2}

convolute on the gray hand image. we can get the gradient directions gx,gy .

Gradient magnitude=$\sqrt{gx2 + gy2}$

Orientation or angle=arctan(gx,gy).

We are eliminating the pixels whose gradient amplitude is less than a mean gradient amplitude multiply by a particular constant(1.2 t0 1.5) and for the remaining pixels LOH is calculated.

$$\sum_{m=1}^{M} \sum_{n=1}^{N} \delta(fgradient\,(m,n) - i),$$

where i is the angle in degrees.

i=arctan(gx,gy).

M,and N image height and widths. $fgradient$ is the matrix of gradient orientations.$\delta(x)$=1 if x=0and zero otherwise. Local Orientation Histogram of static gestures can be seen in [7].

The main idea of using the orientation histogram is that for each gesture there will be a particular orientation histogram. If there is a small change in the hand posture it gives different orientation histogram and The calculation of local orientation histogram is easy and fast.

B. *Gesture Recognition*

After finding out the local orientation histogram for a image Euclidean distance measure used for the current and database LOHs.The database having different LOHs for Different static gestures. The minimum Euclidean distance between the current and database is the matched LOH for the current static gesture. The higher number of training samples will leads to high recognition rate.

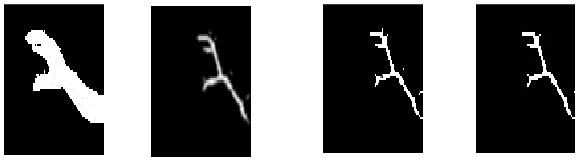
V. DYNAMIC GESTURE RECOGNITION

Skelton concept used for identification of dynamic gestures. Instead of taking several skeleton images taking single image which will covers entire gesture motion called as Dynamic signature.

A. *Skeleton Extraction*

Dynamic signature is the superimposition of all skeleton images in the gesture. This Dynamic signature covers the entire gesture motion. Skeleton means compact representation of an object and preserves the topology of the object. Skeleton is robust against translation rotation and scaling [8].

Skeleton is extracted by using several methods like by using chamfer distance transform [9, 10], Thinning methods [11] (morphological operations). Skeletons are generated by using opencv's library (open source Package that is developed by Intel.) [12].Distance transformation method used for generating skeleton.



a b c d

Figure 3: a. Binary hand image. b. Skelton after applying distance transforms. c.Binary Image of b. d. Noise Removal.

B. Dynamic signature

Dynamic signature means superimposition of all the Skeleton images between the start and stop images. It indicates the entire gesture motion between start and stop images within a single image. Fig 5 indicates the entire gesture motion.

Figure 4: Dynamic signature

This Dynamic signature mainly having three steps 1.preparation (skeletons are overlaid indicates starting of the gesture.) 2. Stroke (skeletons are apart due to the fast movement of the hand). 3. Resting position (Skeletons are again overlapped).

This Dynamic gesture recognition time mainly depends on the position of the gesture in the image. For improving the robustness we created the extreme left and right positions for gesture. We consider the entire motion between these extreme left and right positions.

C .Dynamic Recognition Process

The recognition performed by comparing the obtained dynamic signature with the database images. The comparison is done by using distance measures. The following distance measures are used for recognition. Euclidean distance measure, Image Euclidean distance measure [15, 16], hausdroff distance [17, 18, 19], Chamfer Distance, baddeleys distance [13, 14].

As in the field of computer vision the most commonly used distance is Euclidean distance. It will discard the image structures and cannot represent the actual relationship between the images. If there is a small variation in the image it will rise to the large Euclidean distance value between two images.

Hausdroff distance[17] based comparison between two images can be done as mentioned below. $A=\{a_1,a_2,---a_{mn}\}$ and $B=\{b_1,b_2,---b_{mn}\}$ where A and B are intensity values of A and B images. Where m and n are image width and height.

The Hausdroff distance[17] between two images A and B is defined as

$$H(A,B)=Max(h(A,B),h(B,A)).$$

Where $h(A,B)=\max_{a€A} \min_{b€B} \|a-b\|$ where $\|.\|$ indicates the Euclidean norm between two points a and b.It identifies the point a €A that is farthest from any point b€B and measures the distance from a to nearest neighbor in B. The Hausdroff distance H(A,B) measures the mismatch between two sets. The computation time of hausdroff distance is O(mn) where m and n are image dimensions

VI. COMPUTATION TIME

The computation time for the dynamic signature is higher because of the skeleton computation, which requires several analyses of the image pixels.

The computation of the local orientation histogram easier and takes less processing time because only one time scanning of the image is sufficient for calculation of LOH.

VII. CONCLUSION

Gesture recognition can be done by considering both static gestures and dynamic gestures (Dynamic signature i.e. overlapping of skeleton images).

Static gesture recognition can be done based on freeman –both LOH method [6]for identification of start and stop gestures.LOH calculation Is fast, easy one and robust to illumination changes. Little training is required. The current gesture orientation histogram is compared with the database orientation histograms by using Euclidean distance measure .The best match will have less Euclidean distance value.

The Dynamic signature [7] means the superimposition of all skeleton images. It covers the entire gesture motion in a single image. It reflects the motion and spatial location of the gesture. The current dynamic signature is compared with database dynamic signatures by using image Euclidean distance [15, 16]. The best match having less IMED value. The algorithm is performed on 8 gestures and the recognition rate is 90%in simple conditions (uniform background, limited alphabet).

VIII.REFERENCES

[1] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human-computer interaction: a review,"IEEE Trans. Pattern Anal. Machine Intell., vol. 19, no. 7,pp. 677–695, 1997.

[2] H.-S. Yoon, J. Soh, Y. J. Bae, and H. S. Yang, "Hand gesture recognition using combined features of location, angle and velocity," Pattern Recognition, vol. 34, no. 7, pp. 1491–1501, 2001.

[3] S. Nayaga, S. Seki, and R. Oka, "A theoretical consideration of pattern space trajectory for gesture spotting recognition," in Proc. 2nd IEEE International Conference on Automatic Face and Gesture Recognition (FGR '96), pp. 72–77, Killington, Vt, USA, October 1996.

[4] B. Raytchev, O. Hasegawa, and N. Otsu, "User-independent online gesture recognition by relative motion extraction," Pattern Recognition Letters, vol. 21, no. 1, pp. 69–82, 2000.

[5].F. Quek, "Unencumbered gestural interaction," IEEE Multimedia, vol. 4 , no. 3, pp. 36–47, 1996.

[6]. W. T. Freeman and M. Roth, "Orientation histograms for hand gesture recognition," Tech. Rep. TR-94-03a, Mitsubishi Electric Research Laboratories, Cambridge Research centre, Cambridge, Mass, USA, 1995.

[7]. Bogdan lonescu, Didier Coquin Patrick Lambert Vasile Buzuloiu " Dynamic Hand gesture recognition using the skeleton of the hand"EURASIP journal on Applied signal processing 2005:13, 2101–2109.

[8] C. Arcelli and G. Sanniti di Baja, "Euclidean skeleton via centre-of-maximal disc extraction," Image and Vision Computing, vol. 11, no. 3, pp. 163–173, 1993.

[9]. G. Borgefors, "Distance transformations in digital images,"Computer Vision, Graphics and Image Processing, vol. 34, no. 3,pp. 344–371, 1986.

[10] Y. Chehadeh, D. Coquin, and H. Bolon, "A skeletonization algorithm using chamfer distance transformation adapted to rectangular grids," in Proc. 13th IEEE International Conference on Pattern Recognition (ICPR'96), vol. 2, pp. 131–135, Vienna, Austria, August 1996.

[11] Jonathan Hasthorpe and Nick Mount "The generation of river channel skeletons from binary images using raster thinning algorithms" School of Geography, University of Nottingham.

[12]http://www.willowgarage.com/pages/software/opencv.

[13] A. J. Baddeley, "An error metric for binary images," in Robust Computer Vision, pp. 59–78,Wichmann, Karlsruhe, Germany, 1992.

[14] D. Coquin and Ph. Bolon, "Applications of Baddeley's distance to dissimilarity measurement between gray scale images,"Pattern Recognition Letters, vol. 22, no. 14, pp. 1483–1502, 2001.

[15] Jing Li, Bao-Liang Lu "An adaptive image Euclidean distance" Department of Computer Science and Engineering, Shanghai Jiao Tong University, 800 Dong Chuan Road, Shanghai 200240, China Pattern Recognition 42 (2009) 349 – 357.

[16] Liwei Wang, Yan Zhang, Jufu Feng "On The Image Euclidean Distance" Center for Information Sciences School of Electronics Engineering and Computer Sciences, Peking University Beijing, 100871, China.

[17] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge, Comparing images using the Hausdorff distance, IEEE Trans. Pattern Anal. Mach. Intell. 15 (9) (1993) 850–863.

[18] E.P. Vivek, N. Sudha, Robust Hausdorff distance measure for face recognition, Pattern Recognition 40 (2) (2007) 431–442.

[19] C.H.T. Yang, S.H. Lai, L.W. Chang, Hybrid image matching combining Hausdorff distance with normalized gradient matching, Pattern Recognition 40 (4) (2007) 1173–1181.