

Human Hand Gesture Recognition Using Motion Orientation Histogram for Interaction of Handicapped Persons with Computer

Maryam Vafadar and Alireza Behrad

Faculty of Engineering, Shahed University
Tehran, Iran
vafadar_ma@yahoo.com, behrad@shahed.ac.ir

Abstract. Some groups of handicapped persons cannot reliably move the mouse and do the necessary operation on it to control the computer. However they can do some 3-d hand motions. Variety of tools has been presented for these users to interact with computer. Hand gesture recognition is one of the proper methods for this purpose. This paper presents a new algorithm for hand gesture recognition. In this algorithm, after constructing motion history image of video frames for each gesture and applying necessary processing on this image, motion orientation histogram vector is extracted. These vectors are then used for the training of Hidden Markov Model and hand gesture recognition. We tested the proposed algorithm with different hand gestures and results showed the correct gesture recognition rate of 90 percent. Comparing the results of proposed method with those of other methods showed that in addition to eliminating traditional problems in this area, recognition rate has been improved up to 4 percent.

Keywords: Hand gesture recognition, Motion history image, Hidden Markov model, Motion orientation histogram.

1 Introduction

Some groups of handicapped persons cannot reliably move their hand motions. Although their mental abilities are similar to others, they can not use mouse to interact with computer. However they can do some 3-D hand motions. New tools are needed for these users to interact with computer without mouse limitations [1]. Sensors and video cameras are two sample tools that have been presented for this purpose. Sensor outputs need some filter stages to eliminate temperature, pressure and other noise effects. In addition to low accuracy, higher cost of these devices put them out of reach for general use. Webcams are cheap and widespread [2]. By applying video image processing, they can be used for gesture recognition and interaction of handicapped persons with computer.

Vision-based hand gesture recognition methods are generally categorized in two groups: feature/model based [3] and appearance/view based [4] methods.

In feature/model based methods, initially, we need to extract model or features from images [5]. Implementing this stage often leads to three problems [6]:

1. High computational approximations increase the error rate.
2. High computational complexity slows down the system and makes it unsuitable for real-time interactions.
3. General hand gesture recognition without special features in image frames is not possible.

In appearance/view based methods, images can be used directly for hand gesture recognition. So unlike feature/model based methods, we do not need feature extraction in initial stage [6] and problems of this stage eliminated automatically.

Different appearance/view based methods have been proposed for the task of hand gesture recognition including applying Fourier transform [7], wavelet transform [8] or Principal Component Analysis (PCA) [9-11] on images. Edge orientation histogram also is used for static hand gesture recognition [12]. Using temporal templates is another appearance/view based method for gesture recognition [13-15]. In this method, the basic idea is the projection of the temporal pattern of motion into a single image-based representation called Temporal Template. Then appropriate features are extracted from this image. Temporal Templates have been used for face gesture recognition [13] and hand gesture recognition [14-15].

Several methods have used motion history images (MHI) as temporal templates for gesture recognition [16-19]. In these methods a sequence of image motions are used to constitute a single static image. Then appropriate features are extracted from this image and different methods are used for classification such as least Mahalanobis distance [18], least Euclidian distance [19], Recurrent Neural Network (RNN) [15] and Hidden Markov Model (HMM) [20].

This paper presents a new algorithm for hand gesture recognition. In this algorithm, after constructing motion history image of video frames for each gesture and applying necessary processing on it, motion orientation histogram vectors are extracted. These vectors are then used for the training of Hidden Markov Model and hand gesture recognition. We tested the proposed algorithm with the collected data set and results showed the correct gesture recognition rate of 90 percent.

The paper continues as follows: In section 2, we describe the collected data set of hand gestures. Section 3 explains the method of constructing MHI (Motion History Image). Section 4 presents the extraction of motion orientation histogram from MHI. In section 5, gesture classification using HMM (Hidden Markov Model) is presented. Section 6 shows experimental results for the proposed method and conclusions appears in section 7.

2 Collected Data Set

Data set is collected from 5 persons with different hand size and skin color. No obvious feature or sign exists on the hands of these persons. The collected data set contains different videos of six hand gestures. These gestures are 6 different motions of the hand including right, left, up, downward motion and closing and opening of the

fingers, which is denoted as a, b, c, d, e and f gestures, respectively. Gestures a-d may simulate mouse right, left, up and downward motions and two other motions can be used for the simulation of mouse click. We used 5 different persons for tests and gesture was captured in different conditions. Each person repeated these gestures 10 times in different illumination condition and with different backgrounds. Since handicapped persons may not move their hands exactly in one plane, some of these gestures are not intentionally in a plane. Resolution of the video frames is 176*144.

3 Constructing MHI

3.1 Motion History Image (MHI)

Motion History Image (MHI) is a single static image which its pixel intensity is the function of the motion history of captured sequence at that point. MHI is constructed by successively layering the selected image regions over time. Equation (1) presents constructing MHI.

$$H_{\tau}(x, y, t) = \begin{cases} \tau & \text{if } D(x, y, t) = 1 \\ \max(0, H_{\tau}(x, y, t-1) - 1) & \text{otherwise} \end{cases} \quad (1)$$

where H_{τ} is Motion History Image(MHI) and D is the binary difference image that we will explain it in section 3.2. x and y are image pixels coordinates and t is temporal index. τ , is a threshold for extraction moving patterns in video image sequence. Thus, MHI is a scalar-valued image where more recently moving pixels are brighter.

3.2 Difference Image

Depending on MHI applications, different methods may be used for constructing difference image. One method is the construction of background image and subtraction of the frames from background image [19]. In our application, we may have no information about the background [21]; therefore we used the difference of consecutive frames as difference image after applying a threshold.

Figure 1 shows an example of Motion History Image constructed from difference image. As it is obvious from this figure, more recently moving pixels are brighter than other pixels. Indeed, MHIs show the direction of motion, and we can use this feature for gesture recognition in next sections.

3.3 Noise Filtering

Any noise in video frames is directly transferred to difference images and appears in MHI. This may degrade the performance of gesture recognition algorithm; therefore it is necessary to handle noise.

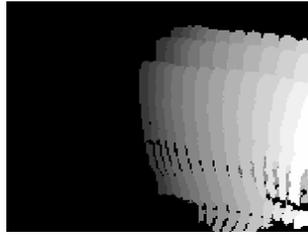


Fig. 1. An example Motion History Image for rightward hand motion

We used morphology operators to remove noise. We applied proper morphology operators to binary image to remove connected component with the size of smaller than 20 pixels. Our experimental results showed that $P=20$ is an optimized value which removes the noise effects and preserves motion information properly. Figure 2 shows the result of noise removing algorithm.

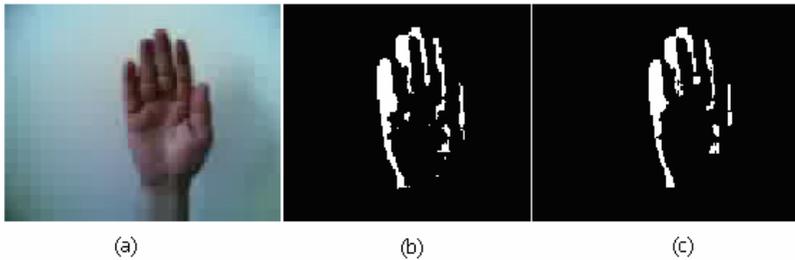


Fig. 2. (a) First frame, (b) difference image and (c) filtered image

3.4 Hand Size Scaling

Proposed algorithm must be insensitive to hand sizes and several users with different hand sizes may use this method. Different hand sizes-even for a single gesture- may have different Motion History Images and results in erroneous results which is not desirable. We applied a normalization algorithm to handle this problem. In normalization methods, hands with different sizes are projected to a constant size. Different algorithm may be used for normalization. In [5] special point are extracted from input image, and distance and angle of the selected points are calculated with respect to image center. These parameters are then used to calculate scaling factor and normalization. Our proposed algorithm doesn't need to extract special points of the image. In this method, we extract the contour of hand motion in MHI using edge detection algorithm. Then rectangular bounding box of the contour is used to constitute an image containing only moving part of the MHI. The resultant image is then resized to standard size of 88×72 . This image is used for hand gesture recognition in next sections. Figure 3 shows the result of normalization algorithm.

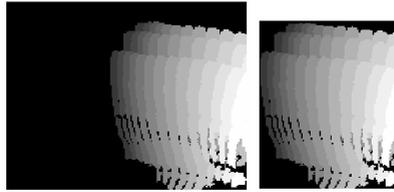


Fig. 3. Hand size normalization of MHI

4 Motion Orientation Histogram

According to equation (1) hand motion results in intensity variation in MHI. So using gradient operator, we can extract motion vector in each pixel and calculate motion fields. We applied Sobel operator to resultant image of previous stage to calculate motion fields as follows:

$$I_x = I * G_x . \tag{2}$$

$$I_y = I * G_y . \tag{3}$$

$$\theta(i, j) = \tan^{-1} \frac{I_y(i, j)}{I_x(i, j)} . \tag{4}$$

where I is Motion History Image of previous stages, G_x and G_y , are Sobel gradient operators. I_x and I_y , are image gradients in x and y directions, showing motion fields and θ is motion orientation and i, j are image pixel coordinates.

When the motion fields are calculated, they are used to extract proper features for classification and gesture recognition. We use the histogram of motion orientations as features for classification and gesture recognition. The features are insensitive to the velocity of motion; therefore they can be used for recognition of the same gesture with different velocities. Figure 4 shows an example of hand motion orientation histogram. To calculate motion orientation histogram, motion orientation is calculated

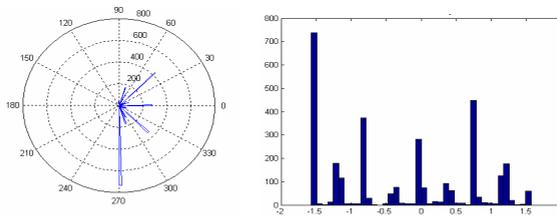


Fig. 4. The orientation histogram for rightward hand motion

using equation (4), then orientation image is divided to 40 bins in polar coordinate, then orientation of points in each bin is summed to obtain orientation histogram. So motion orientation histograms are one dimensional vector of length 40, which are used for classification and gesture recognition.

5 Classification

We use Hidden Markov Model to classify 6 hand gestures. We used left to right model of HMM for classification. This model is appropriate for dynamic gesture recognition. We use One HMM for each hand gesture. In each HMM, motion orientation histogram vectors determines system states. Using training set of motion orientation histogram and baum-welch algorithm, probability parameters are calculated. For hand gesture recognition, after extracting motion orientation histogram, output probabilities $P(O|\lambda^g)$ for all trained models are computed by Viterbi algorithm. At last, a model with highest probability is selected as output.

It is important to note that, we extract orientation histogram for all video samples of a gesture and train all of them to HMM model of that gesture. Weighted Gaussian functions are generated for each gesture by applying k-means algorithm.

6 Experimental Results

We tested the proposed algorithm with our collected data set. We used different combination of gestures as training sets and test inputs. Figure 5 shows hand motion orientation histogram for 6 different gestures. As it is shown in this figure, histograms are different.

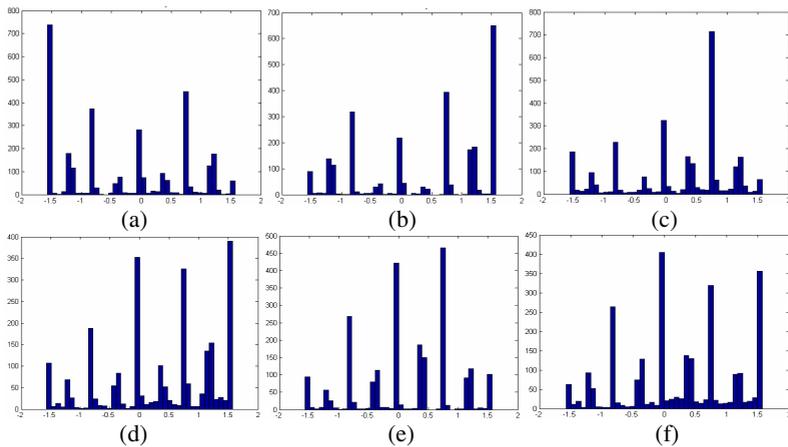


Fig. 5. Motion orientation histograms of (a) rightward (b) leftward (c) upward (d) downward (e) closing of fingers and (f) opening of fingers

We totally conducted 50 different experiments for testing the proposed algorithm and comparing the results with those of other methods. Table 1 shows the results of gesture recognition algorithm for 50 experiments.

To compare the results of proposed method with other methods, we also implemented the algorithms of [18] and [19]. Table 2 shows recognition rates for the proposed and methods of [18] and [19]. Results show the correct gesture recognition rate of 90 percent for the proposed method. Comparing the results of proposed method with two other methods shows that the recognition rate has been improved up to 10 and 4 percent with respect other methods.

Table 1. The result of gesture recognition using the proposed algorithm

<i>f</i>	<i>e</i>	<i>d</i>	<i>c</i>	<i>b</i>	<i>a</i>	Input gestures Algorithm recognition
1	0	0	0	0	47	a
0	1	1	1	48	0	b
0	0	0	45	1	1	c
3	4	44	0	0	0	d
2	42	3	1	1	1	e
44	3	2	3	0	1	f

Table 2. Hand gesture recognition results

method	First method [18]	Second method [19]	Proposed method
Recognition Rate	80,3%	86%	90%

7 Conclusion

A new algorithm based on Motion History Image for hand gesture recognition is presented. In this algorithm, after constructing motion history image of video frames for each gesture and applying necessary processing on this image, motion orientation histogram vector is extracted. These vectors are then used for training of Hidden Markov Model and hand gesture recognition. We tested the proposed algorithm with the collected data set and results showed the correct gesture recognition rate of 90 percent. Comparing the results of proposed method with those of other methods showed that recognition rate has been improved up to 4 percent. In addition the

proposed algorithms is robust to traditional problems of gesture recognition like illumination variations, different skin color and hand sizes. The algorithm can be used for interaction handicapped persons with computer and increase their abilities.

References

1. Derpanis., G.: A Review of Vision-Based Hand Gestures. York University (2004)
2. Hai., W., Shamaie., A., Sutherland., A.: Real Time Gesture Recognition of Human Hand <http://www.computing.dcu.ie/~alistair/CA107.ppt>
3. Mu-Chun, S.: A Neural Network-Based Approach to Recognizing 3D Arm Movement, *Biomedical Engineering Application. Basis and Communication* 15(1), 17–26 (2003)
4. Hou., X., Li., S.Z., Zhang., H., Cheng., Q.: Direct Appearance Models. In: *Proc. Of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 828–833 (2001)
5. Shalhaf, R., Vafadoost, M., Shalhaf, A.: Lipreading Using Image Processing for Helping Handicap. In: *Iranian conf. on Biomedical Engineering* (2007)
6. Trucco., E., Verri., A.: *Introductory Techniques for 3-D Computer Vision*. Prentice-Hall, New Jersey (1998)
7. Harding., P.R.G., Ellis., T.: Recognition Hand Gesture Using Fourier Descriptors. In: *Proc. of IEEE Conf. on Pattern Recognition (ICPR 2004)*, vol. 3, pp. 286–289 (2004)
8. Dionisio, C.R.P., Cesar Jr., R.M.: A Project for Hand Gesture Recognition. In: *Proc. of IEEE Symposium on Computer Graphics and Image Processing*, p. 345 (2000)
9. Lamar., M.V., Bhuiyan, M.S., Iwata., A.: Hand Gesture Recognition Analysis and an Improved CombNET-II. In: *Proc. Of IEEE Conf. on Man and Cybernetics*, vol. 4 (1999)
10. Moghaddam, B., Pentland, A.: Probabilistic Visual Learning for Object Detection. In: *Conf. on Computer Vision*, Cambridge, MA (1995)
11. Moghaddam, B.: Principal Manifolds and Bayesian Subspaces for Visual Recognition. In: *Proc. Of IEEE Conf. on Computer Vision, ICCV 1999* (1999)
12. Freeman, W.T., Roth, M.: Orientation Histogram for Hand Gesture Recognition. In: *IEEE Int. Workshop on Automatic face and gesture recognition* (1995)
13. Black, M., Yacoob, Y.: Tracking and Recognizing Rigid and Non-rigid Facial Motion Using Local Parametric Models of Image Motion. In: *Proc. Of IEEE Conf. on Computer Vision, ICCV 1995* (1995)
14. Shan., C., Yucheng., W., Xianchao., Q., Tieniu., T.: Gesture Recognition Using Temporal Template Based Trajectories. In: *Int. Conf. on Pattern Recognition*, vol. 3, pp. 954–957 (2004)
15. Kumar, S., Kumar, D.K., Sharma, A., McLachlan, N.: Classification of Hand Movements Using Motion Templates and Geometrical based Moments, 3, pp. 299–304 (2004)
16. Venkatesh Babu, R., Ramakrishnan, K.R.: Recognition of Human Actions Using Motion History Information Extracted from the Compressed Video, vol. 22, pp. 597–607 (2004)
17. Ogata, T., Tan, J.K., Ishikawa, S.: High-Speed Human Motion Recognition Based on a Motion History Image and an Eigenspace. *IEICE Trans. on Information and Systems* E89-D, 281–289 (2006)
18. Bobick., A.F.: *Computers Seeing Action*, MIT media library (1997)
19. Davis., J.W.: *Hierarchical Motion History Images for Recognizing Human Motion*, Department of computer and Information science of Ohio state university (2002)
20. Kenny, M., McKenna, S.J.: An Experimental Comparison of Trajectory-Based and History-Based Representation for Gesture. In: *Camurri, A., Volpe, G. (eds.) GW 2003. LNCS (LNAI)*, vol. 2915, pp. 152–163. Springer, Heidelberg (2004)
21. Keskin, C., Erkan, A., Akarun, L.: Real time Hand Tracking and 3D Gesture Recognition for Interactive Interfaces using HMM, Bogazici university, Eusipco (2005)