Hybrid Method for Hand Gesture Recognition Based on Combination of Haar-Like and HOG Features

Sudabeh Ghafouri[1], Hadi Seyedarabi[2]

[1]*University College of Nabi Akram (UCNA), Tabriz, Iran*
[1]`ghafouri_545@yahoo.com`
[2]*Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran*
[2]`seyedarabi@tabrizu.ac.ir`

Abstract— **In this paper a new method is proposed for hand gesture recognition. The proposed method increases hand gesture recognition rate and decreases false positive error rate by using combination of Haar-like and Histogram of Oriented Gradients (HOG) features. Also some new Haar-like features are proposed proportional to hand posture to solve major Haar-like problem that is high false positive error rate in hand posture recognition. These features improve recognition rate to 83%. The experiments showed that hybrid method can recognize hand gesture by 93.5% accuracy which is 25% higher than previous method, and decrease the false positive error from 92% to 8%.**

Keywords— **Hand posture recognition, Haar-like feature, Adaboost learning algorithm, Histogram Gradient Oriented feature, Multi-Class Support Vector Machine**

I. INTRODUCTION

Nowadays with providing new applications of computer such as virtual environment, traditional human-computer interfaces such as keyboards and mice aren't efficient. So, need of natural communication between human and computer is felt. To achieve natural human-computer interaction, hand gestures are important because, after speaking and writing, hand gestures are major human communication way.

Pervious computer vision approaches for hand gesture recognition include 2 steps: 1. posture detection, 2. Hand gesture recognition. In the first step, hand posture is detected by using features like skin color, hand shape and etc. In the second step gesture is recognized by using classification algorithms such as neural networks and syntactical analysis.

Hand posture detection and gesture recognition is difficult job, due to flexible and complex hand structure and also cluttered background. Many studies in this area have tried to find features that have acceptable recognition rate and are robust.

The skin color and the hand shape are image features that are frequently used for hand posture detection [1]. Nevertheless, color-based algorithms face the difficult task of distinguishing objects such as the human arm and the face, which have similar color with the hand. To solve this problem, users are required to wear long-sleeve shirts, and restrictions are imposed on the colors of other objects in the observed scene. Color-based algorithms are also very sensitive to lighting variations. When the lighting does not meet the special requirements, the color-based algorithms usually fail. For shape-based algorithms, global shape descriptors such as Zernike moments and Fourier descriptors are used to represent the hand shape. Most shape descriptors are pixel based, and the computational cost is usually too high to implement real-time systems. Another disadvantage is the requirement of noise-free image segmentation, which is a difficult task for the usually cluttered background images [1]. Some researchers have used machine learning techniques for hand posture detection. Adaboost machine learning algorithm with extract Haar-like feature has been applied in [1]. Major disadvantage of Haar-like feature is high false positive. Zondag et al. compared the performance of two versions of Adaboost with Haar-like and HOG features [2]. The concluded that Haar-like based detector performs, almost, twice faster. The HOG features have the advantage of having a much smaller feature vector than the Haar-like features, so HOG can be used in conjunction with much larger databases. The HOG features detectors consistently achieve better average false positive rates than Haar-like features detectors.

In this paper a new hybrid method is proposed to improve hand posture recognition. Hybrid method is based on combination of Haar-like and HOG features. Hybrid method increases recognition accuracy without increasing false positive error. Also in this paper we designed new Haar-like feature proportional to hand posture and shape.

The reminder of this paper is structured as follow section 2 describes features of image and classification algorithms that used for hand posture detection and recognition. Section 3 describes hybrid method for hand gesture recognition. Section 4 presents our Conclusions.

II. FEATURES AND CLASSIFIERS

In this section we will describe features of image and classification algorithms that used for hand posture detection and recognition.
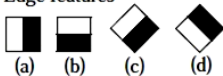
A. features

Haar-like and HOG features have been used for training of Adaboost learning algorithm. Also these features have been used for successfully object detection in [3], [4].

1) Haar-like feature: Haar-like features are rectangular features which first time used for face detection [4]. Then Lienhart and Maydt [5] introduced a novel set of rotated Haar-like features. Fig. 1 shows these extended set of Haar-like features. Each Haar-like feature includes two or three connected "white" and "black" rectangles. The value of Haar-

like feature is the deference between the sum of pixel values in white and black rectangles, i.e, [1]

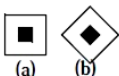$$f(x) = \sum_{black} (pixel\ value) - \sum_{white} (pixel\ value) \qquad (1)$$

Fig. 1 Extended set of Haar-like features [5]

TABLE I
HAND POSTURE RECOGNITION USING HAAR-LIKE FEATURES

	TF	F	P	OF	UK	Recognition Rate
TF	35	1	2	12	0	70%
F	1	20	0	29	0	40%
P	2	3	42	3	0	84%
OF	2	3	5	40	0	80%

Rectangle features can be computed very rapidly using an intermediate representation for the image which was called the integral image. The integral image at location x, y contains the sum of the pixel value above and left of x, y, inclusive:

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x',y') \qquad (2)$$

Using the integral image any rectangular sum can be computed in four array references, as shown Fig. 2 [6].

Chen used Haar-like feature for hand posture detection [1]. These features have high recognition rate for hand posture detection but they suffer from high false positive error rate. Main reason for this problem is that Haar-like features were designed for face recognition which has symmetric structure. Therefore these features are not suitable for hand postures recognition that have complex and asymmetric structures. For Haar-like feature when recognition rate is 99%, false positive is 62.18%, as shown in Table 1. Average recognition rate of Haar-like feature is 68.5%. We tested four hand postures - the "Two Fingers" posture, "Fist" posture, the "Palm" posture and the "One Finger" posture, as shown in Fig. 3. We used Marcel database to train and test the learning algorithm [7].
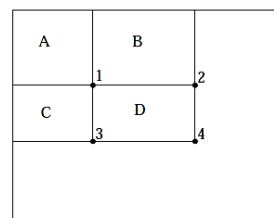

Fig. 2 The sum of the pixels within rectangle D can be computed with for array references. The value of the integral image at location 1 is the sum of the pixels in rectangle A. the value of location 2 is A+B, at location 3 is A+C, and at location 4 is A+B+C+D. The sum within D can be computed as 4+1-(2+3) [6].
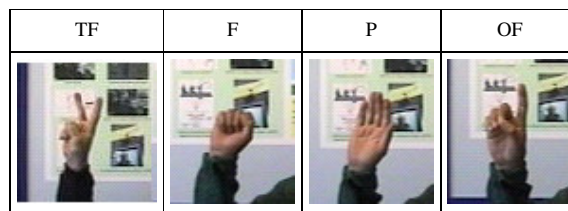

Fig. 3 The tested postures in implementation

To solve Haar-like false positive problem in hand posture recognition, we designed new Haar like feature proportional to hand posture, and added these features to previous Haar-like feature set. Fig. 4 shows new set of Haar-like feature. We did not use integral image to compute new Haar-like feature value because new Haar-like feature are asymmetric. To accomplish that, we assigned zero for black pixel and 1 for white pixel in new Haar-like mask. Then we convolved mask and sub-window of picture. The result of this convolution is sum of pixel values in the white area. Computing sum of pixel values in the black area is done conversely. Then we computed difference between black and White areas. We applied some heuristic methods for increasing speed of computing and decreasing number of false recognitions; For example, we scanned new Haar-like of fist in 1/3 middle of sub-window. New Haar-like features of palm were scanned just in first 1/3 of sub-windows. By Using new set of Haar-like false positive is 38.1% when recognition rate is 99%.

TABLE II
HAND POSTURE RECOGNITION USING NEW SET OF HAAR-LIKE FEATURES

	TF	F	P	OF	UK	Recognition Rate
TF	39	0	2	9	0	78%
F	3	40	0	7	0	80%
P	3	0	45	2	0	90%
OF	2	2	4	42	0	84%

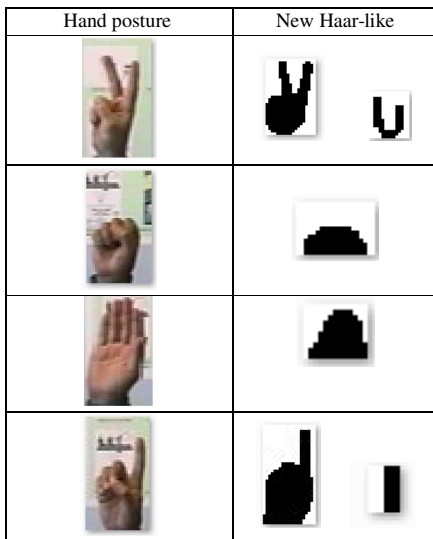| Hand posture	New Haar-like

Fig. 4 new Haar-like feature

The results of hand posture recognition by using a new set of Haar-like features is shown in Table 2. New set of Haar-like feature consist basic Haar-like and new Haar-like features. By adding new Haar-like feature average recognition rate increases to 83%. Fig. 5 compares the average recognition rate of basic Haar-like features and new set of Haar-likes features.
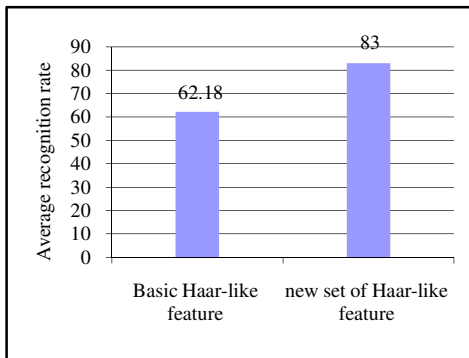

Fig. 5 Average recognition rate of basic and new set of Haar-likes

2) HOG: The histogram of oriented gradients (HOG) [4] encodes the spatial distribution of local intensity gradients. A hand might be well detectable by a characteristic local distribution of edges or intensity gradients. The HOG features are computed by dividing an image into small spatial regions called cells. For each cell a local 1-d histogram of gradient directions is accumulated over the pixels of the cell. The concatenated histogram entries of the cells form the HOG feature vector representation. The gradient directions in the input image are computed, using discrete derivative masks like sobel masks. Each cell level histogram divides the gradient angle range into a fixed number of predetermined orientation bins. Each pixel in the cell votes (weighted) for an edge orientation, based on the orientation of the gradient element centered on it, into the orientation bins (angle ranges) of the cell's histogram. In practice for invariance to illumination changes, a number of cells are combined to form a block and the cells in each block. The Entire feature vector is then a concatenation of all block histograms [2].

B. Classifiers

In this paper we applied Adaboost for hand posture detection and binary recognition. For final posture recognition we applied multi-class SVM[1].

1) Adaboost: Adaboost algorithm is an easy and powerful learning algorithm. Adaboost learning algorithm is based on set of weak classifiers. Accuracy of Weak classifiers is slightly better than random classifiers. Adaboost learning algorithm reweighted the training samples in each step. It increases the weight of samples that have been missed in the previous step of classification. Adaboost combines weak classifiers as linear combination. The final result is a strong classifier. An attentional cascade that was proposed by Viola and Jones [47] is used to increase the speed of process.

2) Multi-class SVM: The Support vector machine was described in detail by Vapnik [8]. SVMs are classifiers which find best separator between two classes. SVMs transfer data into high dimensional space by using kernel function, then it classifies data accurately. Hand posture recognition uses more than two classes, so we use multi-class SVMs to classify hand postures. In this paper we compute the similarities between input image and classes that have been identified by Adaboost. The input image is classified in the class with the highest similarity. There are 4 classes related to the 4 hand postures in our system.

III. HYBRID METHOD

The Inspired by two described features, a hybrid method of hand gesture recognition is proposed. Hybrid method includes combination of Haar-like and HOG features. Hybrid method benefits advantages of Haar-like and HOG features. Therefore it can increase recognition rate while false positive rate remains low.

Table III

HAND GESTURE RECOGNITION USING HYBRID METHOD

	TF	F	P	OF	Uk	Recognition rate
TF	48	0	0	1	0	96%
F	0	48	0	2	0	96%
P	2	0	47	1	0	94%
OF	5	0	1	42	0	88%

The flow diagram is as Fig.6. The main content of hybrid method is as follows. It performs the hand gesture recognition in two steps. First step is done by cascade-Adaboost. Adaboost has been trained with Haar like and HOG features. An image is inputted simultaneously into all hand posture cascade-Adaboost for hand posture recognition. An input image is recognized as a posture when its sub-window has passed from the Adaboost cascade-Adaboost of corresponding posture. For final hand gesture recognition, sub-windows that

---

[1] Support Vector Machine

has been detected as posture, are inputted to multi-class SVM in second step. Multi-class SVM was trained with HOG feature. Result of hand posture recognition by hybrid method is shown in Table 3.
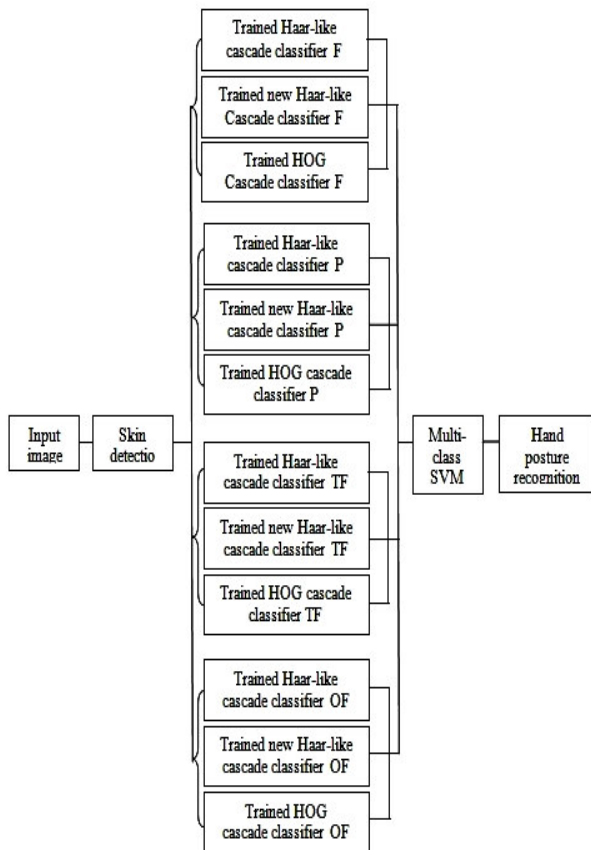

Fig. 6 Hybrid method flow diagram

Fig. 7, Fig. 8 compares average recognition rate and average false positive error rate of above-mentioned methods, respectively. Recognition rate and false positive of hybrid method are 93.5%. Therefore hybrid method is Accurate and reliable.
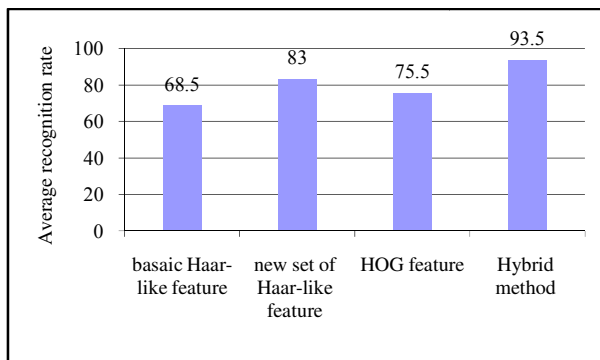

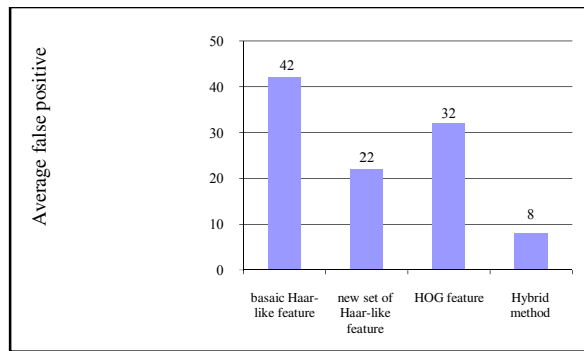Fig. 7 Average recognition rate of mentioned methods

Fig. 8 Average false positive error rate of mentioned methods

IV. CONCLUSIONS

In this paper a hybrid method based on Haar-like and HOG features was proposed. Also some new Haar-like features were designed proportional to hand postures. The proposed approach increases detection and recognition rate while decreases false positive rate. Hybrid method can recognize hand posture in high accuracy. It is better than methods which use Haar-like or HOG features. Hybrid method increases detection rate from 68.5% to 93.5%.

For the future work, there are some suggestions to improve and extend this work. First suggestion is recognition of multiple objects such as eye gaze, human face and hand gesture at the same time. With multi-object recognition strategy, richer commands can be covered. Multi communication makes human-computer interaction more advanced and interesting.

REFERENCES

[1] Q. Chen, N. D. Georganas, and E. M. Petriu, "Hand gesture recognition using haar-like features and a stochastic context-free grammar," *IEEE Transactions on Instrumentation and Measurement*, Vol. 57, pp. 1562–1571, Aug. 2008.

[2] J. A. Zondag, T. Gritti, V. Jeanne, "Practical study on real-time hand detection," *3rd International Conference on ACII*, 2009, p. 1–8.

[3] P. Viola, M. Jones, "Robust reatime object detection," Univ. of Cambridge, CRL Tech. Rep 01–01, 2001.

[4] N. Dala, B. Triggs, "Histograms of oriented gradients for human detection," *In Proc. CVPR'05*, 2005, pp. 886–893.

[5] R. Lienhart, J. Maydt, "An extended set of haar-like features for rapid object detection", *IEEE International Conference on Image Processing*, 2002, pp. 900–903.

[6] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", *IEEE Conference on CVPR 2001*, 2001, pp. 511–518.

[7] Sebastien Marcel Static Hand Posture Database. [Online]. Available: http://www.idiap.ch/resource/gestures/

[8] C. Cortes, V. N. Vapnik. "Support-vector networks," *Machine Learning*, Vol. 20, pp. 273-297, Sept ,1995.