Available online at http://scik.org
J. Math. Comput. Sci. 6 (2016), No. 6, 1002-1011
ISSN: 1927-5307

FINGER ESTIMATION METHOD FOR HAND GESTURE RECOGNITION

ISACK BULUGU*, ZHONGFU YE, JAMAL BANZI

Department of Electronic engineering and Information Science, University of Science and Technology of china, Hefei 230026, China

Abstract. This paper proposes a method for estimating the number of finger from sequences of image frames based on Higher-order Local Auto-Correlation (HLAC) features and multiple regression analysis(MRA). This method is based on a low computation that enables fast and automatic finger extraction for hand gesture recognition in real time. The hand palm region is detected using background subtraction method. Then the palm and fingers are segmented. HLAC features are extracted from image, and multiple regression approach is adopted to estimate number of fingers. Furthermore, in order to improve the performance of the proposed method, we have also proposed a modified version of multiple regression method, and conduct comparative experiment with a normal multiple regression analysis method. The experimental results were analysed, and show that the regularized MRA slightly improves the performance of MRA by stabilizing the results of the estimation.

Keywords: higher-order local auto-correlation; multiple regression analysis; hand gesture recognition.

2010 AMS Subject Classification: 68U05.

1. Introduction

Counting and extraction of fingers in images is very important in several applications, partcularly in hand gesture recognition. Hand gesture recognition play big role in many applications

*Corresponding author

Received July 5, 2016

especially sign language recognition [1] [2] [3] [4], and other similar applications like sign language interpreter for deaf people [5] and robot control [6] [7].The are different researches which are related to the task of counting in images have involved mainly on the detection of a target [8] [9]. In the detection part, detectors are created based on the model of predefined target pattern [10] or using statistical learning [9]. In detection, feature extraction normally works under a small region called detection window,e.g HOG [9] and SVM [11], classifier of the extracted features. The detection window is sliding over the whole image, and it responds only at the position where target is found, and number of object will be counted. Practically, the whole process is exhaustive and difficult since it involves large number of images. In the context of this work, we are extracting HLAC image features effectively from the entire image, even a simple statistical approach gives the best results. In [12], Kurita and Hayamizu proposed a gesture recognition method which uses higher order local autocorrelation (HLAC) features extracted from PARCOR images. Author in [13] propose scale invariant static hand postures detection method using extended HLAC features extracted from Log-Polar images. In this paper, the number of fingers is estimated by means of linear regression (multiple regression analysis), which is applied to HLAC feature vectors extracted from the given image. The contributions of the proposed method is listed as follows; First, we propose a low computation, statistical approach for estimating number of fingers based on Higher-order Local Auto-Correlation (HLAC) and multiple regression analysis for gesture recognition. Second, we modified MRA in order to improve the performance of MRA by stabilizing the results of the estimation. Third, we use a normal and inexpensive camera to capture the vision information.

2. Proposed method

The fundamental idea of a proposed method of hand gesture recognition lies on the result of finger segmentation, and then followed by feature extraction. An overview of the proposed system is shown in figure1 below.

Hand image

⇓

hand detection

⇓

Finger segmentation

⇓

Feature extraction(Binary HLAC)

⇓

Linear regression(MRA,Modified-MRA)

⇓

Finger counting(Estimation)

FIGURE 1. The overview of a proposed method

2.1 Hand detection

We apply the approach in [14] for detecting hand images. The image sequences are captured by a normal camera under the same condition. Since the images have a static identical background, it is easy and effective to apply background subtraction method for the detection of a hand region. The skin colour is measured by HSV model (Hue=315, Saturation=94, and Value=37) more details is found in [15]. The size of the image of the detected hand is made more appropriate so that the gesture recognition is invariant to image scale.

2.2 Finger and palm segmentation

The output of the hand detection is a binary image where by the white pixels represent the members of the hand region, while the black pixels represent the background. Then, the next step is finger segmentation form the binary image by using palm mask. The algorithm 1 is used for searching the palm mask. In order to obtain accurate palm mask a larger circle is used instead of the maximal inner circle.

2.3 Locating hand center

The hand center is regarded as the palm center and insensitive to widely different hand gestures. It can be defined as the maximum distance of a point from the closest hands boundary within the hand region. In this paper, the palm center is found using distance transform method. After the image is transformed to distance map, each pixels distance from the nearest boundary pixel is recorded. Thus, the pixel with largest distance is selected as the hand center point.

2.4 Internal Circle

After the center point is obtained, the internal circle of the maximum radius r is drawn. The radius is extended until it reaches the edge of the palm. That means only white pixels are included inside the circle.

2.5 Wrist Points

In order to obtain wrist points another blue coloured circle of radius 1.2r is produced. Below two below equations represent uniformly sampled points along the circle.

$$X = (Rcos\frac{\theta*\pi}{180})+_{X_o}, Y = (Rcos\frac{\theta*\pi}{180})+_{Y_o}\ \theta = 0 : t : 360$$

Wrist points can be described as two ending points of the wrist line across the bottom of the hand. These wrist points are essential for a successful hand recognition. They can be obtained based on the following condition; Two successive mask points $1(X_i,Y_i), 2(X_i+1,Y_i+1)$ should be large enough compare to others, these two points are described as the wrist points.

$$argmaxdist(P1,P2), P1P2S.$$

Algorithm 1

Input: Uniformly sampled points from the circle.

Output: Boundary points yield the palm mask that can be used to segment fingers and the palm.

Find the nearest boundary point of one sampled (X,Y)

Step 1. Acquire a pixel (x,y) around the sample point (X,Y)

$$x = (cos\frac{\theta*\pi}{180})*rad+X, y = (cos\frac{\theta*\pi}{180})*rad+Y$$

[0, 360] and rad [1,] (is the image size)

Step 2. If the value of the pixel is 0 i.e. $(x,y) == 0$, go to Step 3. Otherwise, increase rad and angle with a step of 1 and then go to Step 1

Step 3. Check the values of 8 neighbors of the pixel (x,y), if it holds $\forall P\ (x+dx, y+dy) == 0, (x+dx, y+dy)N(x,y)N(x,)$is the set of 8 neighbors of the pixel (x,y) .

Step 4. Increase rad and angle, and then go to Step 1 (i) Continue to search the nearest boundary point of another sampled point until all the sampled points are scanned. (ii) Connect all the points recorded in the array of palm mask points to yield the palm mask.

3. Feature extraction

We use the following autocorrelation formula for feature extraction based on Higher-order Local Auto-Correlation (HLAC) [16]. The Nth order HLAC is calculated by following autocorrelations;

$$\int_D (r)I(r+a_1)(r+a_2)....dr \qquad (1)$$

Where $I(r)$ denotes the intensity at the observing pixel r, and 1, 2,.. are n displacements. Since we are dealing with binary images i.e$I(r) = 0,1$, then mask patterns are reduced to 25 mask patterns(shown in figure 2) .The HLAC feature for the binary image is a 25-dimensional vector, called Binary HLAC. It extracts morphological characteristics in the image and has a linear relationship with the Euler number. Each mask pattern is scanned over the entire image, and for each possible position, the product of the pixels marked in white is computed. All the products corresponding to a mask are then summed so as to provide one feature. This operation is performed using 25 different mask pattern to create the feature vector for each hand image.

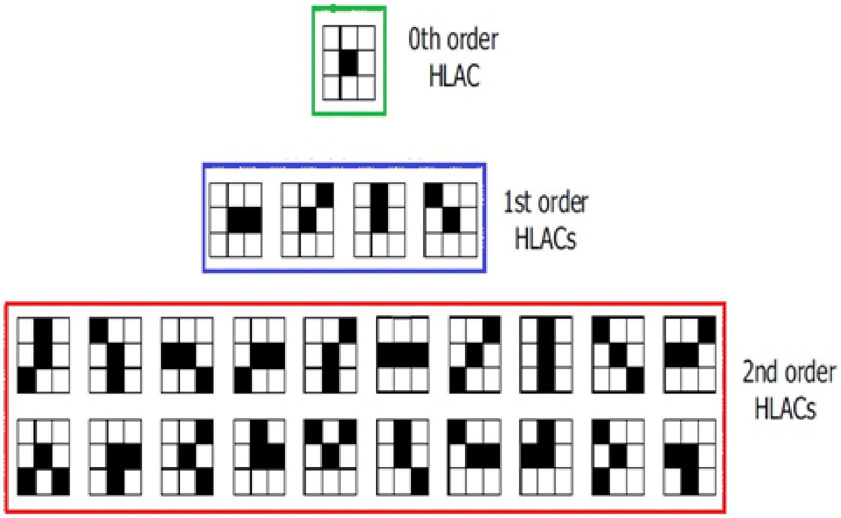

FIGURE 2. Mask pattern

The HLAC-based features have the properties which make counting segmented fingers possible. They are shift invariance and additive to data. Due to the properties of the HLAC-based features, the proposed approach performs hand gesture recognition that is shift invariant to data and additive for data. These properties are categorized as linear methods, and, in particular, are suitable for the task of counting object by eliminating the need for detecting and extracting objects from cluttered backgrounds. According to additive property, suppose that regions A and B are disjoint $(A \cap B = \Phi)$ in an image. Then, the feature vector from the entire image is given as

$$\sum_{r \in A \cup B} \approx \sum_{r \in A} g(r) + \sum_{r \in B} g(r) = R_A + R_B \qquad (2)$$

Where$g(r) = I(r)I(r+a1)I(r+aN)$. This holds because auto-correlations are almost limited to each region $(AorB)$ due to their locality. This property also makes it possible to simultaneously identify multiple objects [15].

4. Multiple Regression Analysis

In the training phase, the pairs of the feature vectors x_i and the correct numbers of segmented fingers C_i for the ith image are given. Then, we determine the optimal linear coefficients$a$, to estimate C from X by using Multiple Regression Analysis (MRA).

$C \approx a^{'} x + b =$â$'$ x̂

where b is constant â=[á+b]',x̂=[x',1]', â can easily be obtained by minimizing regression error

$$\frac{\min}{a,b} \sum_i \parallel c_i - a^{'} xi - b \parallel^2$$
$$\frac{\min}{a,b} \sum_i \parallel c_i - \hat{a}^{'} \ \hat{x} \parallel^2$$
$$\hat{a} = (\ \hat{X}\ \hat{X}^{'})^{-1}\ \hat{X}c \qquad (3)$$

Where X̂= [$\hat{x}_1$,$\hat{x}_2$.....$\hat{x}_N$] and c = [$c_1 c_2$.....c]. Given the image feature x extracted from the test image, the number of segmented fingers in the image can be estimated by For Modified MRA

$$\frac{\min}{a,b} \sum_i \parallel c_i - a^{'} xi - b \parallel^2 - \lambda\ \parallel \acute{a}(x_i - x_{i+1})\ \parallel^{-1}$$
$$\hat{a} = (1+2\lambda)\hat{X}\hat{X}^{'} - \lambda \sum_i \parallel \acute{a}\ (\ \hat{X}\ _i\hat{X}^{'}_{i+1} - \hat{X}\ _{i+1}\hat{X}^{'})\ \parallel^{-1} \qquad (4)$$

The above derived equations prove that the property of additivity in HLAC, the MRA can be applied to the estimation of the number of fingers.

TABLE 1. Configurations of the proposed methods

Method	Feature extraction	Analysis
Binary MRA	Binary HLAC	MRA
Binary MRA	Binary HLAC	R-MRA

Based on flow diagram described in Figure.1, we have two choices, as described above. The configurations of the two proposed methods are shown in Table 1. These methods have the same feature extraction methods for the two types of MRA.

5. Main results

The proposed finger extraction method is used in real-time for hand gesture recognition using MRA and Modified-MRA. The experiments are running with a PC of the following specs; Intel Core i5-4590 CPU running at 3.3 GHz and web camera (f: 4.8). The proposed method is

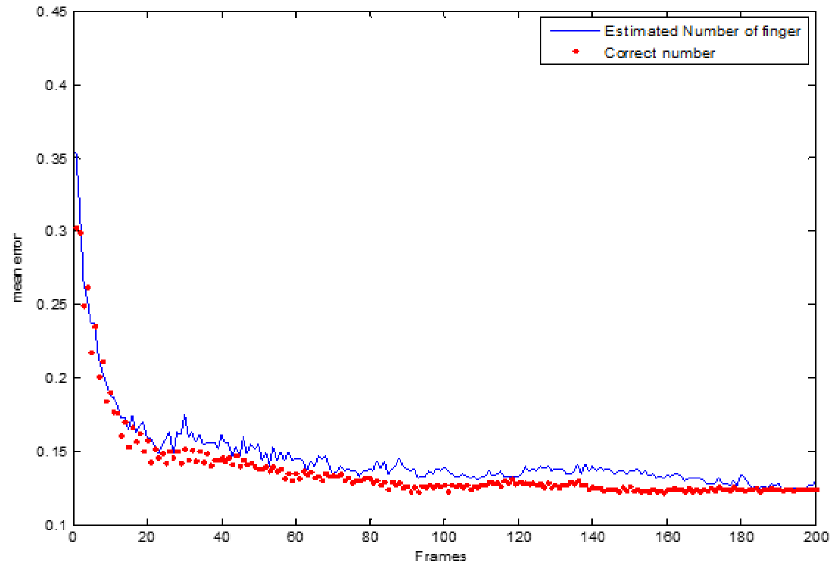

FIGURE 3. MRA

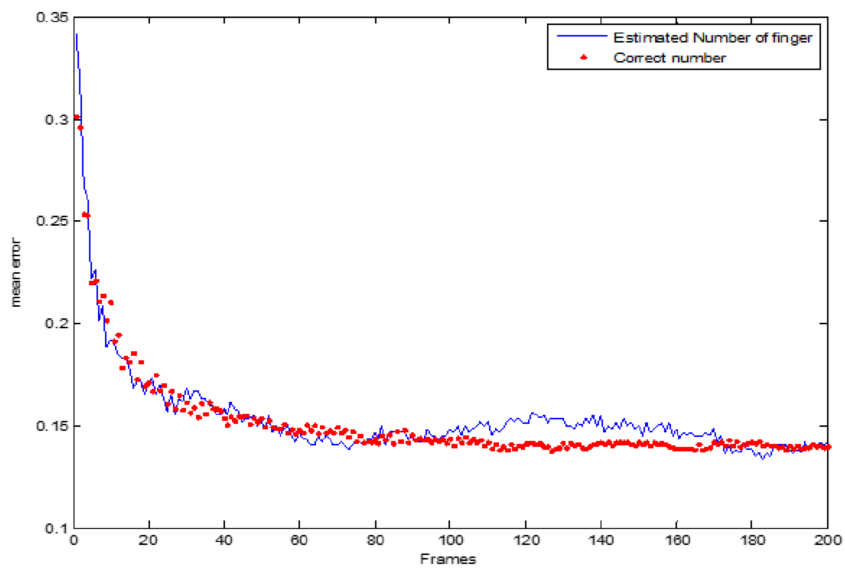

FIGURE 4. Regularized MRA

tested in the experiments, counting segmented fingers for number sign recognition. The images were successively captured at approximately 12fps by monitoring hand movements, and image sequences were obtained. The number of fingers is estimated at each frame. The following condition were considered without loss of generality; the hand should be less 1m distance from the camera, hand palm should be upright facing the camera without finger crossing. The proposed method for finger estimation is tested in the experiment. In order to evaluate the performances of the proposed methods, a leave-one-out cross validation method is applied to these sequences. The linear coefficients are obtained by MRA applied to four sequences and the methods are tested on the remaining sequence, and the mean error of the number of fingers is then calculated. By comparing two figures(3 and 4),the Regularized MRA in figure 4 slightly improves the performance from 0 to 100(frames) whereby the estimated number is nearly similar to correct number eventually stabilizing estimation result.

5. Conclusion

A new method for hand gesture recognition is introduced in this paper. The hand palm region is detected using background subtraction method. Then the palm and fingers are segmented to make easier finger recognition. Furthermore, the fingers in the image are obtained and recognized. The recognition of hand gestures is accomplished by extraction of HLAC image features, and Multiple Regression Analysis (MRA) to estimate number of finger. The properties of additivity and shift-invariance in HLAC are well fit in to the linear model of MRA, and thus require low computation. Modified MRA slightly improves the performance of MRA by stabilizing the results of the estimation. The proposed method can make accurately extraction of fingers and estimation of number of fingers.

Conflict of Interests .

The authors declare that there is no conflict of interests.

REFERENCES

[1] R. Yang, S. Sarkar, B. Loeding, Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32 (2010), 462-477.

[2] Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, P.Presti, American sign language recognition with the kinect, Proceedings of the 13th ACM International Conference on Multimodal Interfaces, 2011, 279-286.

[3] D.Uebersax, J. Gall,M.Van denBergh, L. VanGool,Real time sign language letter and word recognition from depth data,Proceedings of the IEEE International Conference on Computer VisionWorkshops, 2011, 383-390.

[4] N. Pugeault, R. Bowden, ”Spelling it out: real-time ASL fingerspelling recognition,” Proceedings of the IEEE International Conference on Computer VisionWorkshops , 2011, 1114-1119.

[5] J. Zeng, Y. Sun, F. Wang, A natural hand gesture system for intelligent human-computer interaction and medical assistance, Proceedings of the 3rd Global Congress on Intelligent Systems, 2012, 382-385, .

[6] D. Droeschel, J. Stuckler, S. Behnke, Learning to interpret pointing gestures with a time-of-flight camera, Proceedings of the 6th ACM/IEEE International Conference onHuman-Robot Interaction, 2011, 481-488, .

[7] K. Hu, S. Canavan, L. Yin, Hand pointing estimation for human computer interaction based on two orthogonal-views, Proceedings of the 20th International Conference on Pattern Recognition, 2010, 3760-3763.

[8] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection., IEEE Conference on Computer Vision and Pattern Recognition, 2005, 20-25.

[9] P. Viola, M. Jones., Rapid object detection using a boosted cascade of simple features, IEEE Conference on Computer Vision and Pattern Recognition, 2001, 511-518.

[10] T. Shimada, K. Kato, A. Kamikouchi, K. Ito., Analysis of the distribution of the brain cells of the fruit fly by an automatic cell counting algorithm.,Physica A, 350 (2004), 144-149.

[11] V. Vapnik, Statistical Learning Theory, Wiley, 1998.

[12] S. T.Kurita, Gesture Recognition using HLAC Features of PARCOR Images and HMM based Recognizer, Proceedings of The Third International Conference on Automatic Face and Gesture Recognition, 1998, 422 - 427.

[13] Z. Ye. I. Bulugu,Scale Invariant static hand-postures detection using Extended Higher-order Local Autocorrelation features, International Journal of Computer Applications, 135 (2016), 1-5.

[14] Zhi-hua Chen, Jung-Tae Kim, Jianning Liang, Jing Zhang, Yu-Bo Yuan, Real-time hand gesture recognition Using finger segmentation, Scientific World Journal, 2014, p. 9.

[15] A. D. Bagdanov, A. Del Bimbo, L. Seidenari, L. Usai, Real-time hand status recognition from RGB-D imagery, Proceedings of the 21st International Conference on Pattern, 2012, 2456-2459.

[16] T. Kobayashi,T. Hosaka , S. Mimura , T. Hayashi , N. Otsu, HLAC Approach to Automatic Object Counting, Bio-inspired Learning and Intelligent Systems for Security, 2008, 40 - 45.