Proceedings of the 26th Annual International Conference of the IEEE EMBS
San Francisco, CA, USA • September 1-5, 2004

Hand region extraction and Gesture recognition from video stream with complex background through entropy analysis

JongShill Lee[1], YoungJoo Lee[1], EungHyuk Lee[2], SeungHong Hong[1]
[1]Department of Electronic Engineering, Inha University, Korea
[2]Department of Electronic Engineering, Korea Polytechnic University, Korea

Abstract—**The hand gesture recognition utilizing image processing relies upon recognition through markers or hand extraction by colors, therefore it is heavily restricted by the colors of clothes or skin.**

In this paper, we propose a method to recognize hand gestures extracted from images with complex background for more natural interface in HCI(Human Computer Interaction).

The proposed method is obtaining the image through subtract one image from another sequential image, measuring the entropy, separating hand region from images, tracking the hand region and recognizing hand gestures. Through entropy measurement, we've got color information that have near distribution in complexion for region that have big value and extracted hand region from input images. We could draw hand region adaptively in change of lighting or individual's difference because entropy offer color information as well as motion information at the same time. Detected contour using chain code for hand region that is extracted, and present centroidal profile method that is improved little more and recognized gesture of hand.

In the experimental results for 6 kinds of hand gesture, it shows the recognition rate with more than 95% for person and 90~100% for each gesture at 5 frames/sec.

Keywords—**entropy, hand gesture, contour, recognition**

I. Introduction

With the massive influx of computers in society, HCI(Human Computer Interaciton) has become and increasingly important part of our daily lives. In recent years there has been a tremendous push in research toward novel devices and techniques .

One long-term attempt in HCI has been to migrate the "natural" means that humans employ to communicate with each other into HCI. With this motivation automatic speech recognition has been a topic of research for decades. Tremendous progress has been made in speech recognition, and several commercially successful speech interfaces have been deployed[1]. However, it has only been in recent years that there has been an increased interest in trying to introduce other human-to-human communication modalities into HCI. This includes a class of techniques based on the movement of the human arm and hand, or hand gestures.

They range from simple actions of using our hand to point at and move objects around to the more complex ones that express our feelings and allow us to communicate with others. To exploit the use of gestures in HCI it is necessary to provide the means by which they can be interpreted by computers. The HCI interpretation of gestures requires that dynamic and/or static configurations of the human hand, arm, and even other parts of the human body, be measurable by the machine. First attempts to solve this problem resulted in mechanical devices that directly measure hand and/or arm joint angles and spatial position. This group is best represented by the so-called glove-based devices [2], [3], [4], [5], [6]. Glove-based gestural interfaces require the user to wear a cumbersome device, and generally carry a load of cables that connect the device to a computer.

Potentially, any awkwardness in using gloves and other devices can be overcome by using video-based *noncontact* interaction techniques. This approach suggests using a set of video cameras and computer vision techniques to interpret gestures. Many of those approaches have been chosen and implemented so that they focus on one particular aspect of gestures, such as, hand tracking, hand posture estimation, or hand pose classification. Until recently, most of the work on vision-based gestural HCI has been focused on the recognition of static hand gestures or postures.

The hand gesture recognition utilizing image processing relies upon recognition through markers or hand extraction by colors, therefore it is heavily restricted by the colors of clothes or skin. Therefore we propose a new method to recognize hand gestures extracted from images with complex background for more natural interface in HCI.

II. Extraction and tracing of hand region using entropy analysis

The proposed method is obtaining the image through subtract one image from another sequential image, measuring the entropy, separating hand region from images, tracking the hand region and recognizing hand gestures.

The proposed method is obtaining the image through subtract one image from another sequential image, measuring the entropy, separating hand region from images, tracking the hand region and recognizing hand gestures. Through entropy measurement, we've got color information that have near distribution in complexion for region that have big value and extracted hand region from input images.

We could draw hand region adaptively in change of lighting or individual's difference because entropy offer color information as well as motion information at the same time.

Deected contour using chain code for hand region that is extracted, and present centroidal profile method that is improved little more and recognized gesture of hand.

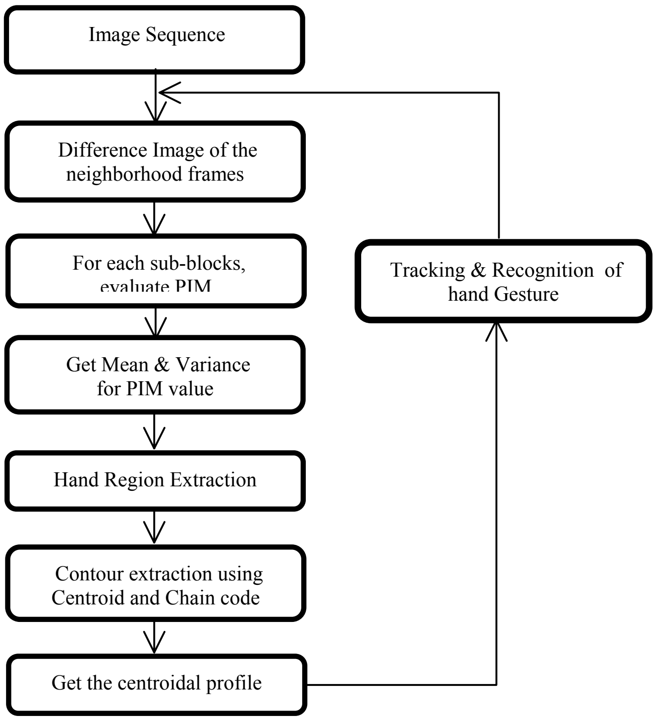

Fig. 1. Block diagram of the proposed algorithm for hand gesture recognition.

A. PIM-based Moving Detection

In general, it is required to measure the information of image when one design a PIS(Pictorial Information System). And, if the quantity of information is small, the information will not stored or stored with compressed. For example, in the case of video communication, image which has small information must will be transmitted with compressed.

Thus, it is important to use the device which measures the information of image. In this paper, entropy which is among the pixel is used to obtain the characteristic of image data, and PIM which was suggested by Chang is introduced to quantify the entropy of image[5]. The PIM method which is used in this paper is given by the following equation.

$$PIM = \sum_{i=0}^{L-1} h(i) - Max_j h(i) \tag{1}$$

In equation, $h(i)$ means i-th histogram value of each image or block. *PIM* value is evaluated by subtracting the total number of pixels in each block form the histogram value of maximum frequency. If all the pixel value of block is same, that is the block entropy is '0', (1) becomes (2), and thus *PIM* value becomes '0'.

$$\sum_{i=0}^{L-1} h(i) = Max_j h(i) \tag{2}$$

If each of the level values of pixels are distributed uniformly in block, that is the case of high entropy, the value of $Max_j h(i)$ is small and it means high *PIM* value.

We can get the maximum *PIM* value as following pseudo-code.

For M×N size, L-gray image
if M×N≤L,
then Max(PIM)=M×N-1
if M×N〉L,
then Max(PIM)=M×N-[M×N/L]
([n] : maximum integer value not exceed n)

According to the pseudo-code, *PIM* value is high when the block has large quantities of information, low when the block has small quantities of information.

And *NPIM*(Normalized *PIM*), PIM_k , $NPIM_k$ is given in (3), (4), (5).

$$PIM = \sum_{i=0}^{L-1} h(i) - Max_j h(i) \tag{3}$$

$$PIM = \sum_{i=0}^{L-1} h(i) - Max_j h(i) \tag{4}$$

$$PIM = \sum_{i=0}^{L-1} h(i) - Max_j h(i) \tag{5}$$

Fig. 2 shows the extraction of hand region in complex background image using *PIM* which use subtraction image.

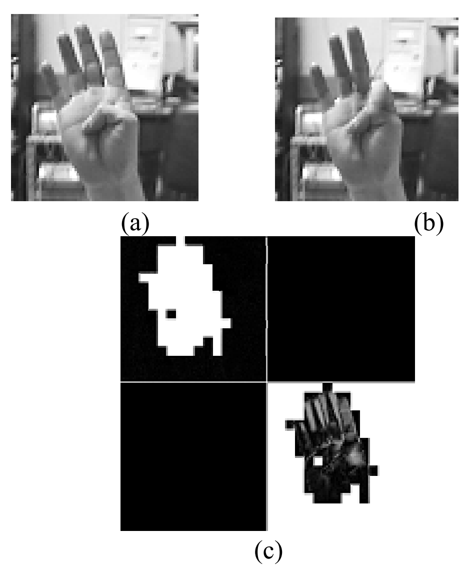

Fig. 2. Extraction of Hand Region using PIM

B. Tracking of Hand Region

In this paper, we apply suggested algorithm which is drawn Fig. 1 to real sequential image for the tracking of hand region and recognition of sign-language.

In Fig. 3, (a) shows input images and (b) shows detected hand region by using suggested algorithm.

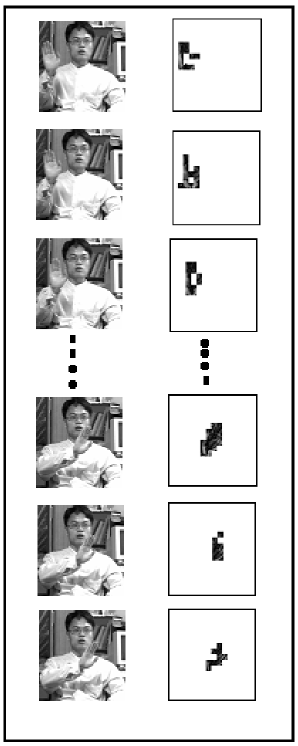

Fig. 3. Tracking of hand region using the proposed method.

C. Extraction of Hand Region

Through entropy measurement, we've got color information that have near distribution in complexion for region that have big value and extracted hand region from input images. We could draw and detect hand region adaptively in change of lighting or individual's difference because entropy offer color information as well as motion information at the same time.

Fig. 4 shows an example of the hand region extraction through entropy analysis.

(a) (b) (c)

Fig. 4. Hand Extraction through entropy analysis.
(a) Given hand gesture, (b) Using fixed color distribution, (c) Using adaptive color distribution through entropy analysis.

D. Recognition of Hand Recognition

There are several problems with the centroidal profile approach. The (r,θ) plot will be multi-valued for a certain class of object. This has the effect of making the matching process partly 2-D and leads to complication and excessive computation. Therefore we propose the modified centroidal profile. First, we get the contour using chain code for the given hand region. And we compute the centroidal profile for each pixel of the contour. That is, we compute the distance from the centroid of the hand region to the contour boundary.

The distance(r) of each point $(\bar{x},\bar{y})$ on the contour from the center of mass can be computed by (6), (7), (8).

$$\bar{x} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} jI(i,j)}{A}, \quad \bar{y} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} iI(i,j)}{A} \tag{6}$$

$$A = \sum_{i=1}^{N}\sum_{j=1}^{M} I(i,j) \tag{7}$$

$$r = \sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \tag{8}$$

III. Results

We performed the experiment for 6 kinds of hand posture which are shown in Fig. 5.

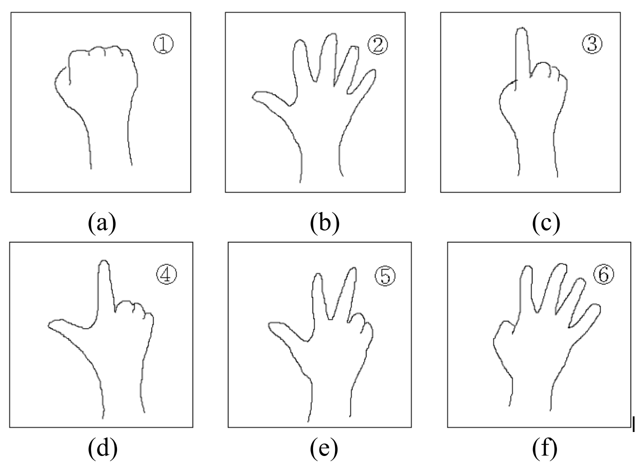

Fig. 5. 6 kinds of hand posture

We used USB camera for the input image device and tested in Pentium 650Mhz.

As shown in Fig. 6, we could extract hand region adaptively in change of lighting or individual's difference. Detected contour using chain code about hand region that is extracted, and presented centroidal profile method that was improved little more and recognized gesture of hand.

The resutlts of hand region extraction and centroidal profile using the proposed method for sample hand posture is shown in Fig. 6.

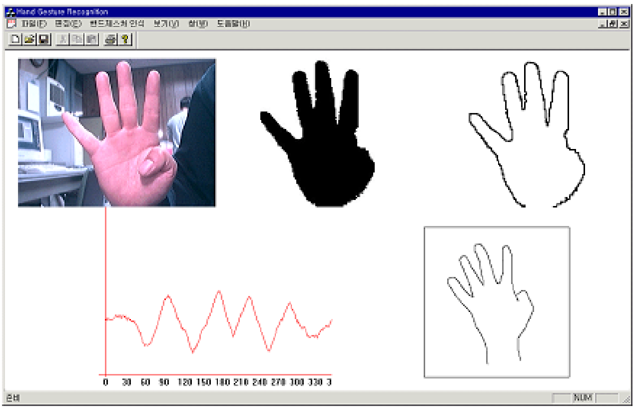

Fig. 6. Hand gesture recognition system.

In the experimental results for 6 kinds of hand gesture, it shows the recognition rate with more than 95% for person and 90 ~ 100% for each gesture as shown in Table I.

TABLE I
Recognition rate of hand posture using the proposed algorithm

Posture / person	1	2	3	4	5	6	Total
#1	20/20 (100%)	19/20 (95%)	20/20 (100%)	18/20 (90%)	19/20 (95%)	20/20 (100%)	116/120 (97%)
#2	20/20 (100%)	20/20 (100%)	20/20 (100%)	19/20 (95%)	20/20 (100%)	19/20 (95%)	118/120 (98%)
#3	20/20 (100%)	18/20 (90%)	20/20 (100%)	19/20 (95%)	19/20 (95%)	18/20 (90%)	114/120 (95%)
#4	20/20 (100%)	20/20 (100%)	20/20 (100%)	18/20 (90%)	19/20 (95%)	19/20 (95%)	116/120 (97%)
#5	20/20 (100%)	18/20 (90%)	20/20 (100%)	19/20 (95%)	20/20 (100%)	19/20 (95%)	116/120 (97%)
#6	20/20 (100%)	19/20 (95%)	20/20 (100%)	20/20 (100%)	20/20 (100%)	20/20 (100%)	119/120 (99%)

IV. Conclusion

The proposed method is obtaining the image through subtract one image from another sequential image, measuring the entropy, separating hand region from images, tracking the hand region and recognizing hand gestures.

Through entropy measurement, we've got color information that have near distribution in complexion for region that have big value and extracted hand region from input images.

We could draw hand region adaptively in change of lighting or individual's difference because entropy offer color information as well as motion information at the same time. Detected contour using chain code for hand region that is extracted, and present centroidal profile method that is improved little more and recognized gesture of hand.

In the experimental results for 6 kinds of hand gesture, it shows the recognition rate with more than 95% for person and 90~100% for each gesture at 5 frames/sec.

References

[1] L.R. Rabiner and B. Juang, Fundamentals of Speech Recognition.Englewood Cliffs, N.J.: Prentice Hall, 1993.
[2] V. I. Pavlovic, R. Sharma and T. S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction : A Review," IEEE Trans. PAMI., vol. 19, no. 7, 1997.
[3] Oh-Young Cho, Hyoung-Gon Kim, Sung-Jea Ko, Sang Chul Ahn, "A Hand Gesture Recognition System for Interactive Virtual Environment," IEEK, 36-s(4), pp. 70-82, 1999.
[4] T. S. Huang and V. I. Pavlovic, "Hand Gesture Modeling, Analysis and Synthesis," Proc. Int'l Workshop on Automatic Face and Gesture Recognition, pp. 73-79, 1995..
[5] F. K. H. Quek, "Toward a vision-based hand gesture interface," Proc. of the Virtual Reality System Technology Conf., pp. 17-29, 1994.
[6] D.J. Sturman and D. Zeltzer, "A Survey of Glove-Based Input," IEEE Computer Graphics and Applications, vol. 14, pp. 30-39, Jan. 1994.