Motion Segmentation using Hand Movement and Hand Shape for Sign Language Recognition

Takanori Shiosaki, Tadashi Matsuo, Yoshiaki Shirai, Nobutaka Shimada
Ritsumeikan University/Department of Human and Computer Intelligence, Shiga, Japan

***Abstract*— For sign language recognition, Hidden Markov Model (HMM) has been used. To build reliable HMMs, each state in a HMM should correspond to a simple motion. To find necessary states, a segmentation method was proposed. However, by the method, we cannot segment motions such as waving a hand, where the center of gravity is almost fixed. For detecting the change of hand shape, we proposes a feature that reflects the difference of appearances. This feature consists of the circularity and the direction of the inertia principal axis. By using the proposed feature, we segment 5 sign language words. The result of the experiments show the effectiveness of the proposed method.**

I. Introduction

Sign language is a communication tool for hearing-impaired people. Because few hearing people can use a sign language, a system is needed to translate the sign language into natural language.

In our method, hand shape features are extracted from video images because use special equipments such as data gloves[1] is troublesome.

Sign language words are trained and recognized using Hidden Markov Model (HMM)[2]. HMM has some states as shown in Fig. 1, where each state has probability distributions of features. Since collecting many samples for sign language recognition is generally difficult, it is desired that each state can be trained with not so many samples. For building reliable HMM, it is desired that each state corresponds to a type of motion such as raising hands, spreading hands, waving a hand, etc.

Kawahigashi et al. proposed a method to segment image sequences for generating necessary states[3]. However, their method is based on the motion of the gravity center of hand regions in image. Accordingly, their method cannot segment a motion where the gravity center is fixed, but the shape varies.

We propose a method to segment a motion with fixed gravity center. This is based on the inertia principal axis and the circularity.

II. Features used in segmentation

We focus on the motion where the contour of hand shape changes. Practically, the inertia principal axis direction varies in such motion even if the gravity center is fixed. To detect the change of the hand shape, we use the inertia principal axis direction θ_p. They are defined as

$$\begin{aligned} I(\theta) &\stackrel{\text{def}}{=} \int_{\text{Hand}} \|\boldsymbol{x}-\boldsymbol{\mu}\|^2 - |\langle \boldsymbol{x}-\boldsymbol{\mu}, \boldsymbol{e}_\theta \rangle|^2 \, dxdy, \\ \theta_p &\stackrel{\text{def}}{=} \arg\max_\theta I(\theta), \end{aligned} \tag{1}$$

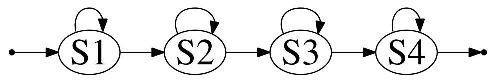
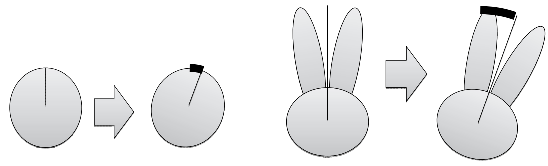
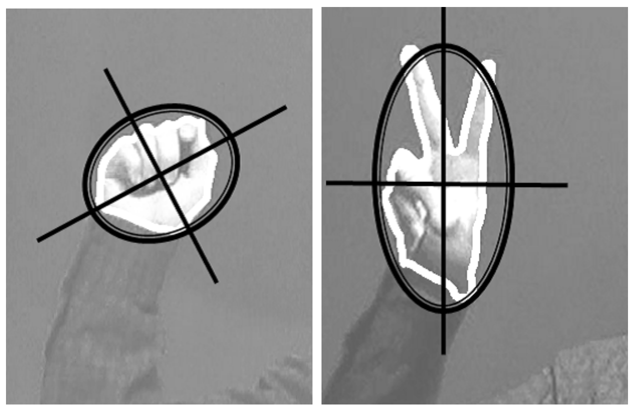


Fig. 1. Hidden Markov Model



(a) a shape like a circle (b) a thin shape

Fig. 2. The circularity and significance of rotation

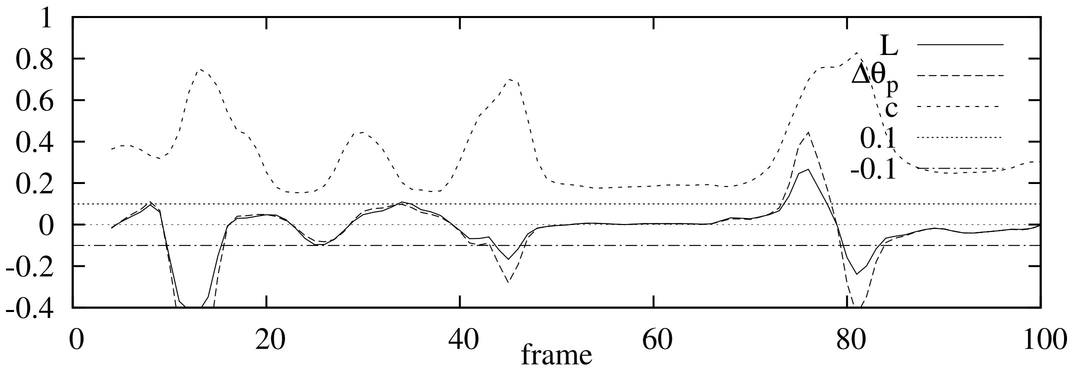


(a) $c = 0.69$ (b) $c = 0.25$

Fig. 3. Example of circularity

where $\|\boldsymbol{x}\|$ is the length of the vector $\boldsymbol{x}$, $\boldsymbol{e}_\theta$ is the unit vector with the direction θ, $\boldsymbol{\mu}$ is the gravity center of the hand region, $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is the inner product of the vectors. By evaluating $\Delta\theta_p$, the temporal difference of θ_p, we can detect the change of the contour.

$\Delta\theta_p$ represents the change of the principal direction of hand shape. However, it is not meaningful when $\min_\theta I(\theta)$ is close to $\min_\theta I(\theta)$ such as shown Fig. 2. This corresponds to the fact that a circle has rotational symmetry. Therefore, we have to change the measure of $\Delta\theta_p$ according to the circularity. To



(a) The transition of L

(b) Segmentation with $\Delta\theta_p$

(c) Segmentation with L

Fig. 4. The segmentation result for the word "NN"

take account of circularity, we define the circularity c as

$$c \stackrel{\text{def}}{=} \frac{\min_\theta I(\theta)}{\max_\theta I(\theta)}. \tag{2}$$

The two example values of c are shown in Fig. 3.

When $\Delta\theta_p$ is large, the motion may be segmented at the time. However, $\Delta\theta_p$ is unreliable if c is close to 1. Accordingly, we introduce a new feature L as

$$L \stackrel{\text{def}}{=} \frac{\Delta\theta_p}{\sqrt{\frac{c}{c_0}}}, \tag{3}$$

where c_0 means the standard circularity(here, we use 0.25). Since $1/\sqrt{c}$ corresponds to the ratio of the major axis to the minor axis, L can be considered as the arc length drawn by the rotation shown in Fig. 2.

We segment the motion at the frame where L exceeds the threshold 0.1. Since we use 30 frames per second, the threshold corresponds to the semicircular rotation of the ellipse whose circularity is c_0.

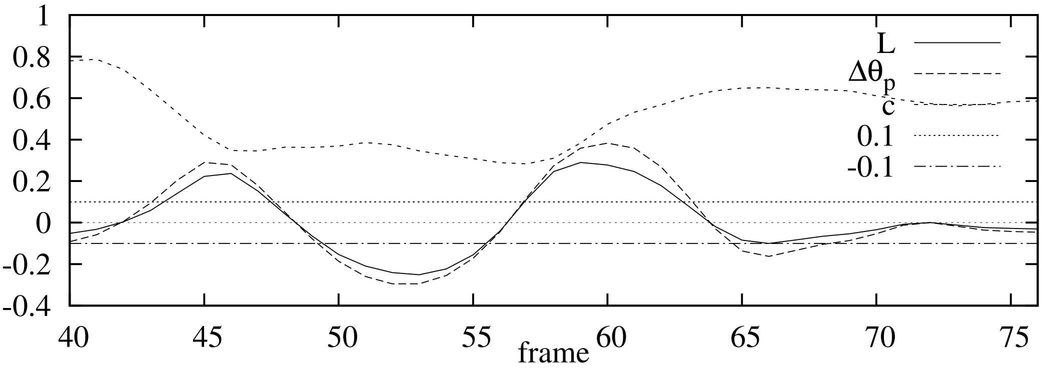
III. Experiment

For segmentation with hand shape we take 5 sign language words, where the gravity center is almost fixed and the hand shape changes. The words are selected from the 6th and 7th grade of sign language proficiency test. For each word, two persons perform the motion 3 times.

We show two results for the word "NN" (a character of Japanese Hiragana) and the word "hot" in Fig. 4, 5. The transitions of L in Fig. 4(a) and Fig. 5(a) show that L is smaller than $\Delta\theta_p$ if $c > 0.25$. For the word "NN", the first segment by $\Delta\theta_p$ (the first column in Fig. 4(b)) includes multiple motions. By using L, the motion is segmented into three segments shown in Fig. 4(c). For the word "hot", the last two segments by $\Delta\theta_p$ (the last two columns in Fig. 5(b)) correspond to rotating motions, where hand shape is similar to circle. By using L, the motions are integrated in a segment.

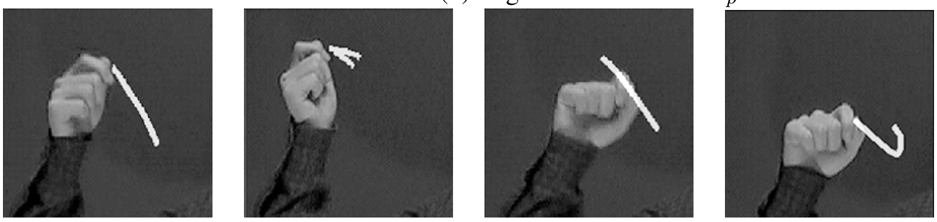
IV. Conclusion

We proposed a feature for detecting changes of motions. It is based on the fact that the rotation of a shape which is similar to a circle is not so significant. The result of the experiments



(a) The transition of L

(b) Segmentation with $\Delta\theta_p$



(c) Segmentation with L

Fig. 5. The segmentation result for the word "hot"

shows that we can automatically segment motions where the gravity center is fixed. By integrating the proposed method, we will improve the method of generating HMMs.

References

[1] H. Sagawa and M. Takeuchi, "A method for recognizing a sequence of sign language words represented in a japanese sign language sentence," in *FG '00: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*. Washington, DC, USA: IEEE Computer Society, 2000, p. 434.

[2] J. Kinscher and H. Trebbe, *Mathematical Background of HMM*, Münster Tagging Project - Arbeitsbereich Linguistik, University of Münster, 2005. [Online]. Available: http://santana.uni-muenster.de/Publications/hmmmath.ps

[3] K. Kawahigashi, Y. Shirai, N. Shimada, and J. Miura, "Segmentation of sign language for making HMM," *IEICE technical report*, vol. 105, no. 67, pp. 55–60, 20050513, (in Japanese). [Online]. Available: http://ci.nii.ac.jp/naid/10016435220/