SICE Journal of Control, Measurement, and System Integration, Vol. 2, No. 5, pp. 263–270, September 2009

Gesture Recognition Based on the Probability Distribution of Arm Trajectories

Khairunizam WAN * and Hideyuki SAWADA *

Abstract : The use of human motions for the interaction between humans and computers is becoming an attractive alternative to verbal media, especially through the visual interpretation of the human body motion. In particular, hand gestures are used as non-verbal media for the humans to communicate with machines that pertain to the use of the human gestures to interact with them. This paper introduces a 3D motion measurement of the human upper body for the purpose of the gesture recognition, which is based on the probability distribution of arm trajectories. In this study, by examining the characteristics of the arm trajectories given by a signer, motion features are selected and classified by using a fuzzy technique. Experimental results show that the use of the features extracted from arm trajectories effectively works on the recognition of dynamic gestures of a human, and gives a good performance to classify various gesture patterns.

Key Words : upper body motion, arm trajectory, fuzzy technique, gesture recognition.

1. Introduction

Humans are able to communicate with each other not only using verbal cues but also non-verbal cues such as gesture, audition and tactility. Through the combination of verbal and non-verbal cues, humans smoothly communicate with each other by expressing their emotions and intensions. A goal in human-computer interaction (HCI) is to improve the interaction between humans and computers by making computers more usable and receptive to user's needs. For achieving the goal, computers should be able to understand human emotions and intensions through body movements and facial expressions that are effectively employed in the transmission of non-verbal information [1]–[4].

In many current human-computer interactions, keyboards, mice and joysticks are used as the user interfaces, and have become common devices for controlling and navigating machines. Such mechanical devices, however, are inconvenient and unsuitable for the natural and direct interactions, since they require to be touched by a user for the operation of the machine, and are impractical in mobile computing applications. A human gesture allows a user to communicate with machines in the natural way, and provides an attractive communication tool that could archive the goal of interacting humans and computers.

Modern three-dimensional tracking systems or motion capture systems allow us to collect features from human movement activities, and a computer is programmed to recognize motions for communicating with a human. Recently, various types of motion capture systems using sensors such as a dataglove, a magnetic position sensor, and a vision sensor have been introduced for the purposes of measuring human motions. Such sensors have a certain advantage in comparison with each other, and are used depending on user applications. A vision-based motion capture system allows a performer to move freely in front of a camera, and has the ability of measuring unlimited movements.

In this paper, we propose a dynamic gesture recognition system using a vision-based human motion capture. An optical motion capture system is used for extracting motion features from arm trajectories, which are the movement and velocity of the arms, and followed by the resampling of motion data in the pre-processing process. A fuzzy technique is used for the classification of the motion features, which is based on the distribution of the resampled motion data. In the recognition process, the aggregation of the fuzzy information is proposed to recognize a certain gesture given by a signer. In the experiments, thirty sign languages are chosen to evaluate the performance of the system.

This paper is structured as follows: Section 2 addresses the related researches to the identifications, approaches and problems of recognizing human gestures. Section 3 describes the configuration of the developed system. Section 4 introduces the proposed methodology used in the study, which presents an important theoretical aspect. Section 5 presents an experiment and its results by using the proposed gesture recognition algorithm. Section 6 discusses the contribution of the works to HCI, and we conclude the article with the summary in section 7.

2. Related Works

Recently, many studies for recognizing human gestures have been reported, which use different parts of the body represented by a human, involving the hands, head and face [2]–[4]. Among them, hand gestures play an important role in the transmission of non-verbal information owing to the flexibility of the human hands. Hand gestures can be described by four attributes, which are hand shapes, palm orientations, hand locations and arm motions [5]. However, many works in the recognition of hand gestures have been concentrated on the classification of the shapes and locations of hands [5]–[9], and relatively a few works have been done on the classification of dynamic arm motions. Arm

* Faculty of Engineering, Kagawa University, 2217-20, Hayashi-cho, Takamatsu-city, Kagawa 761-0369, Japan
E-mail: s06d507@stmail.eng.kagawa-u.ac.jp, sawada@eng.kagawa-u.ac.jp
(Received December 12, 2008)
(Revised March 13, 2009)

motions could be used to differentiate the semantics of hand gestures that have similar shapes and locations of hands. Previously, several approaches to recognize hand gestures based on arm motions were proposed for the implementation in the various applications. J. Lementec et al. used multiple orientation sensors for the recognition of arm gestures to control umanned aerial vehicle (UAVs), and the system was experimented in the laboratory environment for controlling a robot [10]. Hienz et al. developed a video-based system for the acquisition and classification of hand-arm motions concerning a sign language recognition [11]. In the research a geometric model of a human arm was constructed due to the lack of information provided by a single camera.

In the recent years, with the advancement of the computational technologies, various algorithmic techniques for performing accurate recognition of hand gestures are introduced [9],[12]–[14]. F. Bevilacqua et al. applied Hidden Markov Model (HMM) for recognizing hand gesture patterns for the application of gesture interface to support music pedagogy [15]. T. Starner et al. also used HMM for the real time recognition of a sign language using a single camera to track unadorned hands of a user [16]. The interesting points regarding their work are the purpose of recognizing sentence-level of American sign language (ASL) selected from 40-words lexicon consisting of pronouns, nouns, verbs and adjectives. H. Sawada et al. reported a robust gesture recognition using a fuzzy sensor approach for the aggregation of gesture information based on the resampling of data [17]. In this study, by following their proposed approach, an algorithm was developed for recognizing a dynamic hand gesture based on arm trajectories.

In this paper, we introduce a robust algorithm for the recognition of the dynamic hand gestures, which is based on the dynamic movement of arm motions. Recognizing arm motions is a challenging study because of the human variability that causes the same motions into different performances each time due to the inclusions of different emotions [5]. Based on the analysis of gestural motions, we found the importance of emotional information presented in human gestures, although most of the existing works are concentrated on the shapes and locations of hands [12]–[16]. Even for the same person, different instances of the same motions may not identical because of the difference in the speed and size of motions. By segmenting a motion into the same distance-interval data using the resampling algorithm, gestures are translated to be independent of the speed and size of motions. Moreover, experimental results show that for recognizing hand gestures, instead of extracting hand shapes, the information provided by arm trajectories could work well for the classification among various gesture patterns.

3. System Configuration

An optical motion capture system which tracks 7 markers simultaneously in real time is used for the motion measurement. The system is equipped with 2 high-speed cameras with an image resolution of 640×480 pixels and the ability of capturing 120 frames per second. Arm motions are mainly used to analyze hand gestures excluding hand shapes, due to the complexities of the finger configurations and occlusion problem. Seven markers are attached to the body of the signer, which include the chest (Ch), shoulders (R_Sh, L_Sh), elbows (R_El, L_El) and hands (R_Ha, L_Ha) for both right and left upper body. R and L represent the position of the right and left sides of the upper body. Figure 1 shows the flow of the recognition algorithm and it is divided into two phases, which are the learning and recognition phases. In the learning phase, from the captured 3D position data of the reflective markers attached to the body, the system tracks the moving markers and computes the characteristic features of arm trajectories, which is followed by the resampling of motion data. By using the fuzzy technique, the membership functions are designed for the classification of the various gesture patterns based on the distribution of resampled motion data, which are stored in the database as a gesture template. In the recognition phase, after the resampling of data, motion vectors are compared with the features stored in the template. Gesture pattern recognition is executed based on the similarity measured between input motion vectors with the registered gestures.

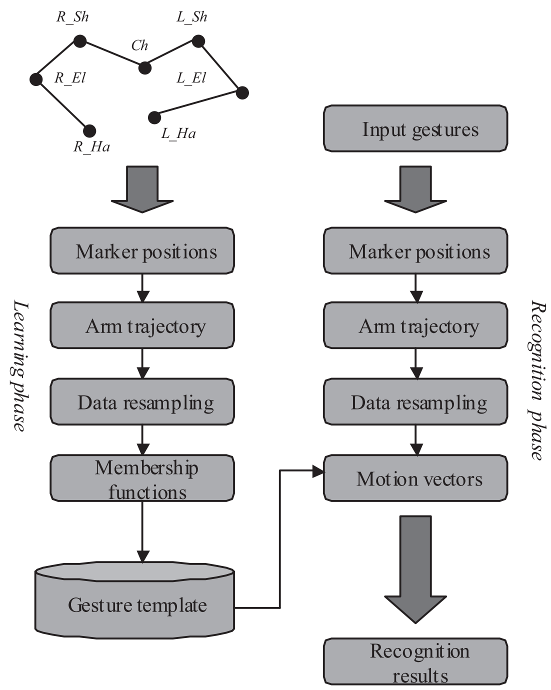


Fig. 1 The flow of the recognition algorithm.

4. Methodology

4.1 Upper Body Model

The outputs from the motion capture system are the 3D positions of 7 markers, and the upper body model is needed for the correspondence with the body parts. The model is constructed based on a kinematics constraint, where the lengths of human limbs are represented using sticks and their relative joint angles stored as illustrated in Fig. 2. The lengths of the sticks $\boldsymbol{T} = \boldsymbol{T_j}$ $(j = 1, 2, 3, \ldots, 6)$ are given based on the earliest frames of a certain pose, and $\boldsymbol{T_j}$ is a translation vector $\boldsymbol{R^3}$ in homogeneous coordinates. Positions of the body parts are estimated and determined by 9 parameters $\boldsymbol{\Theta_i}$ $(i = \gamma, \beta_1, \varepsilon_1, \eta_1, \alpha_1, \beta_2, \varepsilon_2, \eta_2, \alpha_2)$, where $\boldsymbol{\Theta_\gamma}$ is the rotation matrix of the upper body at the chest position. $\boldsymbol{\Theta_\beta}$, $\boldsymbol{\Theta_\varepsilon}$ and $\boldsymbol{\Theta_\eta}$ are the rotation matrices at the shoulder and $\boldsymbol{\Theta_\alpha}$ is the rotation matrix at the elbow. The distances between the markers and the limitation of the joint angles are used to identify body parts in motions. For discriminating body parts using the location data of markers, and by following the representation of affine transformations, the equations

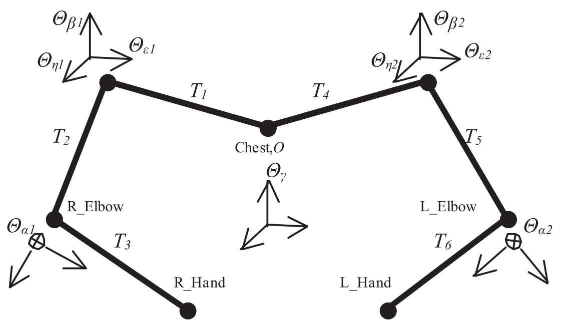


Fig. 2 Skeleton model of the human upper body.

$$\begin{bmatrix} \vec{H}_{RT} \\ 1 \end{bmatrix} = \begin{bmatrix} \Theta_{\alpha 1} & \vec{T}_3 \\ 0,\cdots,0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \Theta_{\beta 1}\Theta_{\varepsilon 1}\Theta_{\eta 1} & \vec{T}_2 \\ 0,\cdots,0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \Theta_{\gamma} & \vec{T}_1 \\ 0,\cdots,0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \vec{O} \\ 1 \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} \vec{H}_{LT} \\ 1 \end{bmatrix} = \begin{bmatrix} \Theta_{\alpha 2} & \vec{T}_6 \\ 0,\cdots,0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \Theta_{\beta 2}\Theta_{\varepsilon 2}\Theta_{\eta 2} & \vec{T}_5 \\ 0,\cdots,0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \Theta_{\gamma} & \vec{T}_4 \\ 0,\cdots,0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \vec{O} \\ 1 \end{bmatrix} \tag{2}$$

are established by referring to Fig. 2. The equations are used for determining the points that are connected by the sticks T_1, T_2, T_3 and T_4, T_5, T_6 of the right and the left arms represented by the transformations $\boldsymbol{H_{RT}} = (x_R, y_R, z_R, 1)^T$ and $\boldsymbol{H_{LT}} = (x_L, y_L, z_L, 1)^T$, respectively [18].

4.2 Arm Trajectories

The objective of the system is to recognize the gestures given by a signer in front of a stereo camera, and to focus on the arm trajectory. Additionally, the movement sequences of both right and left arms are used for the gesture recognition. As the features for recognizing arm motions that represent on behalf of the gestures, the movement and velocity of the hands and elbows are employed for the classification and recognition among various gesture patterns.

4.3 Probability Distribution of Arm Trajectories

Based on the positions of 7 reflective markers, the characteristic features of arm trajectories are computed in the preprocessing process. The trajectories are different with each other even a signer gives the same gesture due to the difference of the speed and size of motions. Furthermore, motion features obtained from arm trajectories are based on time-interval, and the interval distance among sampled points changes according to the difference of motion speed as shown in Fig. 3 (a). To resolve the trajectory difference, a treatment algorithm is needed to resample the time-interval motion features to the distance-interval based data. For obtaining the distance-interval based data, a resampling algorithm was introduced in our previous study [17], and it is applied in this study to do the comparison of data after the resampling process.

Captured position data are divided into 30 equal length pieces. Each point is determined and stored as resampled points. In this manner, the difference of speed and size is absorbed to be compared with the same trajectories. Figure 3 shows the results of the resampling process. Before the resampling, it is difficult to do the comparison of data of the same motions as shown in a circle in the figure. After the resampling, similar motion trajectories are obtained, which have similar displacement intervals, and it is possible to do the comparison between them.

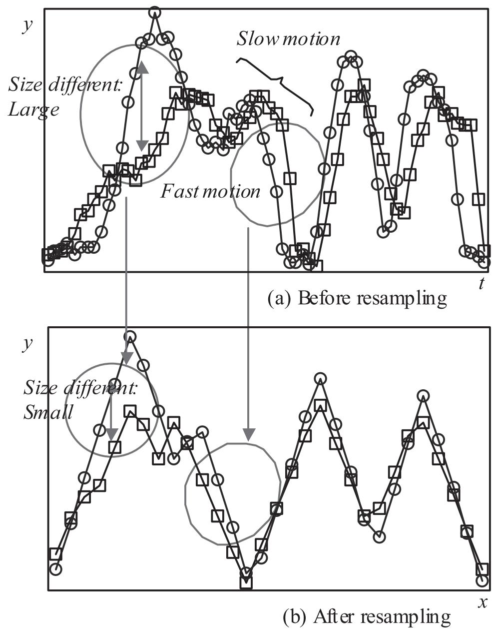


Fig. 3 Resampling results.

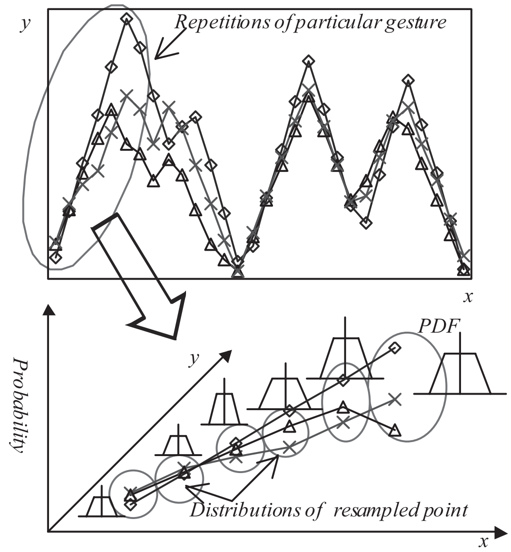


Fig. 4 Distribution of resampled data and the generation of membership functions (MFs).

Figure 4 shows that with the repetitions of a particular gesture, similar trajectories are obtained, even if a signer gives a variation in speed and size. Each resampled datum point is distributed close to each other, and the distribution may give the probability of data derived from a particular gesture. To describe the distribution of the resampled motion data informatively, we got an idea to use the probability density functions (PDFs) for applying fuzzy.

Figure 5 shows the PDF used in the system, and is designed by using the average and dispersion of the distributed data. The PDFs represent the membership functions (MFs) of a fuzzy set, and is generated for each distributed resampled point as shown in Fig. 4. During the learning phase, the averages and standard deviations of the distributed resampled points are stored in the database as the motion parameters that represent the types of gestures. During the recognition phase, the fuzzy membership functions are automatically generated using the stored motion parameters, and output the probability values corresponding to

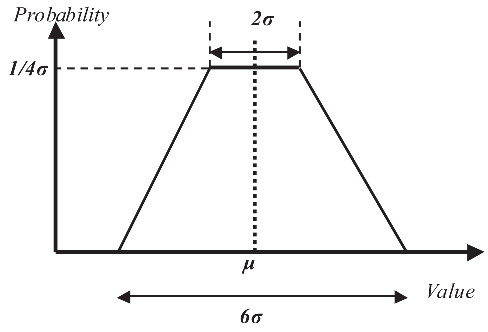


Fig. 5 Probability density function (PDF).

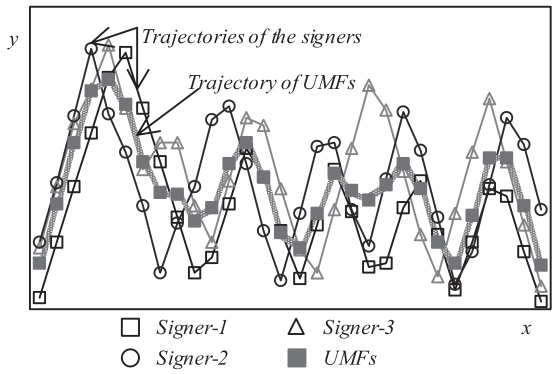


Fig. 6 Trajectories of the distributed resampled data of the signers and UMFs.

the input of fuzzy information. For the recognition of gestures, fuzzy aggregation approach is proposed in this study.

4.4 Universal Membership Functions (UMFs)

The membership functions (MFs) described in the section 4.3 are designed to each signer for the recognition purpose, which means signers have their own templates, where their own gestures are registered. The averages and standard deviations of the distributed resampled data of a particular user are stored as the motion parameters in the database as described in the section 4.3. For designing the universal membership functions (UMFs), the averages and standard deviations of the distributed resampled data are obtained from gestures performed by all signers. Figure 6 shows the trajectory of the distributed resampled data of UMFs, which are measured by averaging the trajectories of the distributed resampled data of the 3 signers. Figure 7 shows the design of the UMFs, and the distributions of its resampled motion data are stored as a new gesture template. During the recognition phase, UMFs are automatically generated using the motion parameters that are stored in the database. UMFs are designed for the purpose of the gesture recognition by sharing the same MFs among all the signers.

4.5 Gesture Recognition

In the recognition part, the system is trained to recognize a gesture performed by a signer, which is based on the arm trajectories that contain fuzzy information. For a given gesture trajectory, the output of the MFs is obtained that represents the type of gestures performed by a signer. The fuzzy aggregation method is employed for the recognition of the gestures, and is based on the intersection operator, which is essential in the fuzzy theory [19]. Let X be a space of points with a generic element of X denoted by x. For the two fuzzy sets A and B of

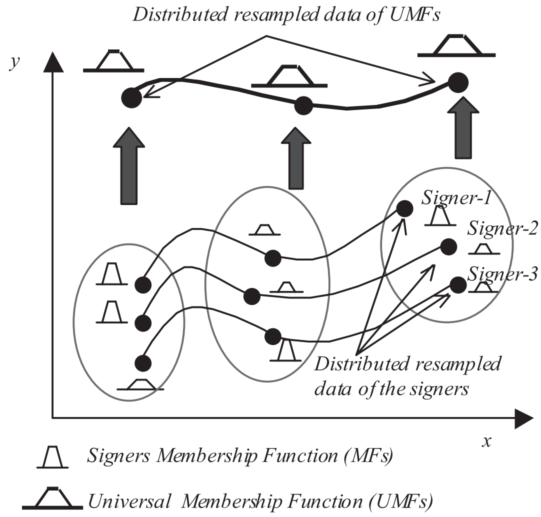


Fig. 7 The design of universal membership function (UMFs).

X, the intersection of $A \cap B$ is

$$A \cap B \rightarrow \mu_{A\cap B}(x) = min[\mu_A(x), \mu_B(x)] \tag{3}$$

$$\mu : \; A \; membership \; function$$

$$x \in X$$

In this study, a measurement is designed to recognize a gesture by using the intersection operation to the two fuzzy sets D_{Right_arm} and D_{Left_arm} of the arm motions as shown in Eq. (4). The fuzzy subset $D_{Gesture}$ is the probability, which represents the gesture obtained from the motion combination of the right and left arms.

$$D_{Gesture}(i) = D_{Right\_arm}(i) \cap D_{Left\_arm}(i) \tag{4}$$

$D_{Gesture}$, D_{Right_arm} and D_{Left_arm} represent the overall meaning of the gesture, gestures of the right and left arms, respectively. (i) is an index of the gestures registered in the database. The equations

$$D_{arm(R,L)}(i) = d_{elbow(R,L)}(i) + d_{hand(R,L)}(i) \tag{5}$$

$$d_{elbow}(i) = \alpha \cdot d^e_{move}(i) + (1-\alpha) \cdot d^e_{velo}(i) \tag{6}$$

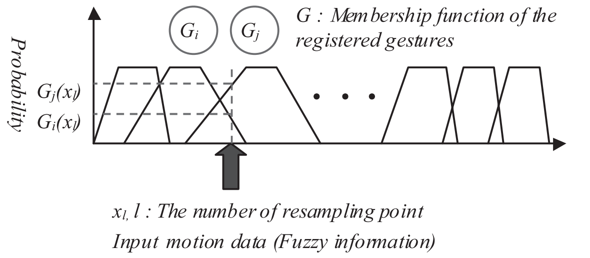
$$d_{hand}(i) = \alpha \cdot d^h_{move}(i) + (1-\alpha) \cdot d^h_{velo}(i) \tag{7}$$

show the formulations for the recognition of the gesture of the right and left arms, denoted by $D_{arm(R,L)}$. α is the weight used in the experiment for obtaining a relevant recognition result. The formulations d_{elbow} and d_{hand} consist of d_{move} and d_{velo}, which are the movement and velocity of a joint, respectively. Each of them consists of 3 components, which are x, y and z components in the 3D space, and is denoted by $\gamma^{[l,a]}$ and $\nu^{[l,a]}$, respectively. γ and ν represent the membership degree of the movement and velocity, where l and a are the number of the resampling point and 3D Cartesian coordinate, respectively. Equations (8) and (9) show the mathematical relations of the movement and velocity of the arms.

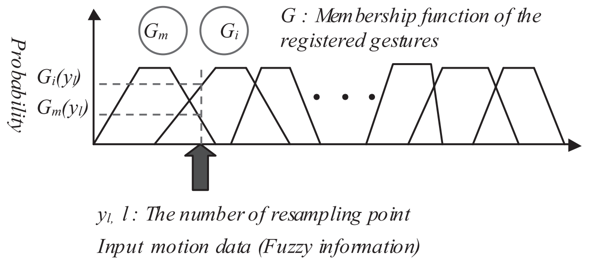
$$d_{move}(i) = \sum_{l=1}^{30} \sum_{a=x,y,z}^{3} \gamma^{[l,a]}(i) \tag{8}$$

$$d_{velo}(i) = \sum_{l=1}^{30} \sum_{a=x,y,z}^{3} \nu^{[l,a]}(i) \tag{9}$$

Figure 8 shows the process of the proposed recognition algorithm. For recognizing gestures, the trajectory sequences given

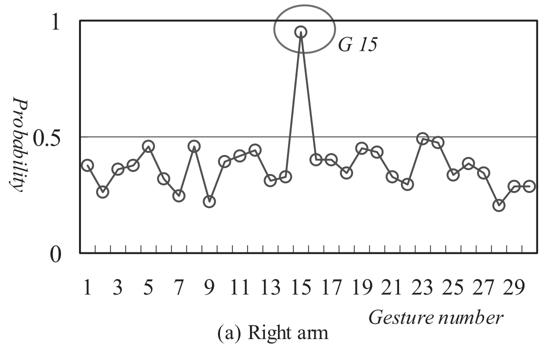


(a) Arm movement d_{move}

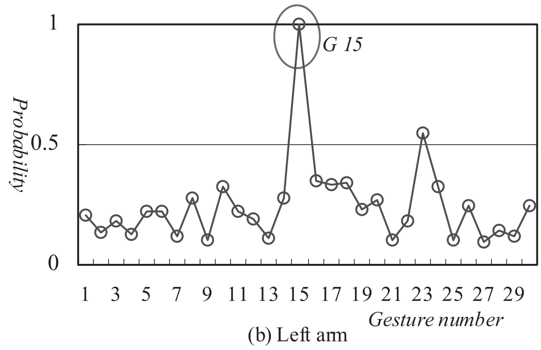


(b) Arm velocity d_{velo}

Fig. 8 Recognition of the gesture.



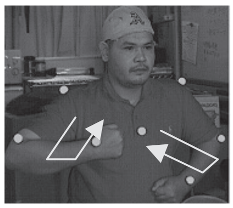
(a) Right arm



(b) Left arm

Fig. 9 Probability of the arm motions.

(a) Rain



(b) Danger

(c) Excercise

Fig. 10 Three examples of the gestures.

by the right and left arms are considered, which employ the movement and velocity of the arms denoted as d_{move} and d_{velo}, respectively. The input motion data to the MFs are x_l and y_l, which are corresponding to the fuzzy information of the movement and velocity of the arms, and l is the number of the resampling point. The MFs give outputs, which are the probability values generated by each resampled point of the input motion data. By referring to the equations (8) and (9), d_{move} and d_{velo} are defined as a probability score obtained by the summation of the probability values outputted from the MFs.

For example, Fig. 8 (a) shows the probabilities $G_j(x_l)$ and $G_i(x_l)$ that are given by the movement of the right arm d_{move}. Fig. 8 (b) presents the probabilities $G_m(y_l)$ and $G_i(y_l)$ that are given by the velocity of the right arm d_{velo}. By employing the fuzzy aggregation, the recognized gesture of the right arm is defined as the intersection of $\{G_j(x_l), G_i(x_l)\}$ and $\{G_m(y_l), G_i(y_l)\}$, which gives the highest probability score D_{Right_arm} where the right arm of a signer performs a gesture G_i. Figure 9 shows the probability of the right and left arm trajectories, which are measured by aggregating values outputted from the MFs. By combining the trajectory sequences of the right and left arms, the system outputs the probability scores $D_{Gesture}$ showing that the overall meaning of the gesture performed by a signer is a gesture "G 15".

5. Experiments

5.1 Experimental Setup

To test and evaluate the proposed methodology in section 4, the experiments were conducted using a large set of sign languages from Japanese Sign Language (JSL) database $\boldsymbol{G}$, where 30 words are registered as

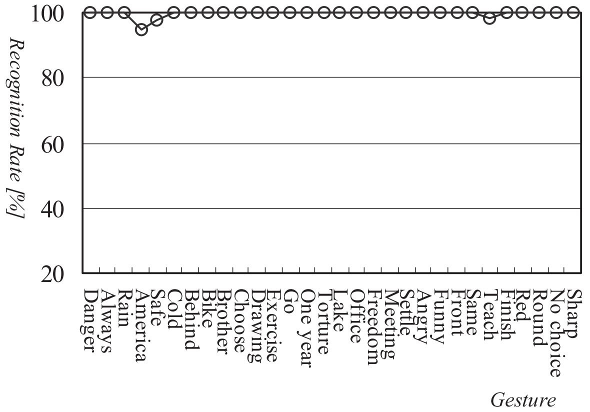
$\boldsymbol{G}$ = {*Danger*, *Always*, *Rain*, *America*, *Safe*, *Cold*, *Behind*, *Bike*, *Brother*, *Choose*, *Drawing*, *Exercise*, *Go*, *One year*, *Torture*, *Lake*, *Office*, *Freedom*, *Meeting*, *Settle*, *Angry*, *Funny*, *Front*, *Same*, *Teach*, *Finish*, *Red*, *Round*, *No choice*, *Sharp*}

The examples of the gestures were shown in Fig. 10. In the learning phase, each gesture was inputted 10 times for the generation of the membership functions (MFs). In the recognition phase, five signers numbered by 1-5 performed all the gestures, and each gesture was given 10 times with the various size and speed in the motion.

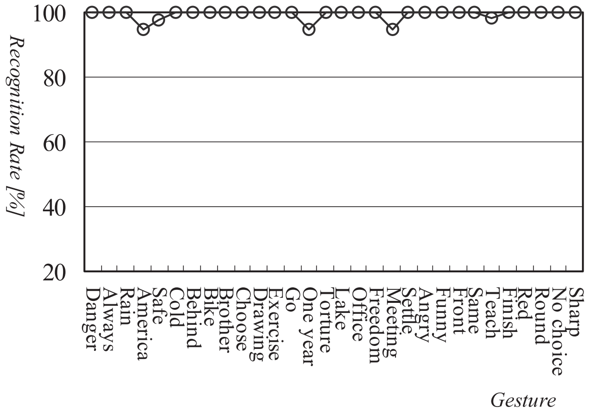
5.2 Experiment Results

5.2.1 Gesture recognition using the individual membership functions (MFs)

In the first experiment, the gestures given by the 5 signers were recognized using their individual MFs. The recognition results show that the system successfully recognized almost 100% of all gestures given by the signers, and Fig. 11 shows the examples of the results of a signer-1 and signer-4. Table 1 shows the accuracies for all 5 signers using their individual MFs. The accuracies shown in the table were computed by averaging the results of all gestures given by the signers in the experiments. The results show that the system had a high per-

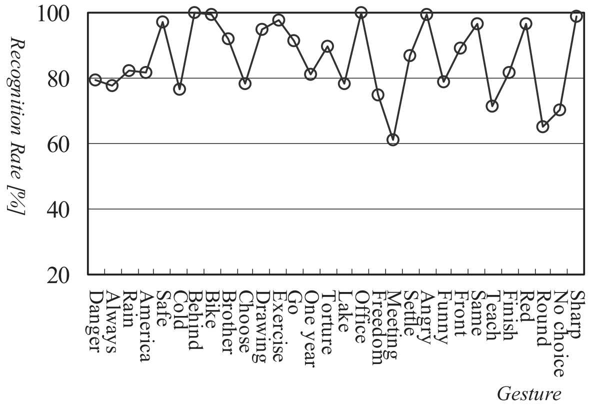


(a) Signer-1

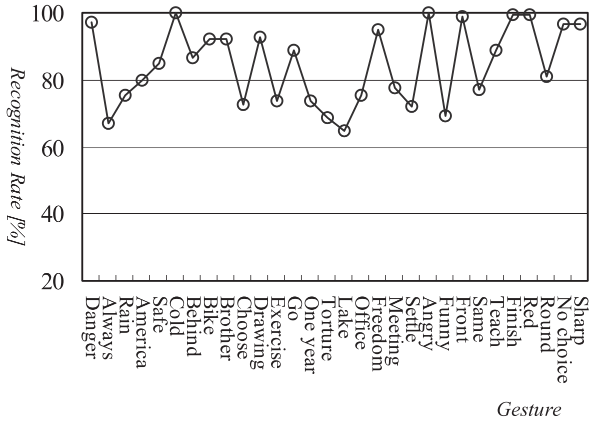


(b) Signer-4

Fig. 11 Recognition results using the individual MFs.



(a) Signer-1



(b) Signer-4

Fig. 12 Recognition results using signer-5's MFs.

Table 1 Accuracies of the signers using their individual MFs.

Signers	*1*	*2*	*3*	*4*	*5*	*AVG*
Recognition rate (%)	99.8	97.9	99.6	99.7	99.0	99.2

Table 2 Accuracies using other signer's MFs.

MFs / *Signers*	*Signers 1*	*Signers 2*	*Signers 3*	*Signers 4*	*Signers 5*	*AVG*
Signer 1	***	88.2	83.6	84.6	87.4	86.0
Signer 2	81.5	***	80.0	82.5	86.7	82.7
Signer 3	87.9	85.9	***	84.6	86.6	86.2
Signer 4	77.5	83.2	76.5	***	84.6	80.5
Signer 5	85.6	86.6	83.4	83.0	***	84.7

formance for recognizing various types of the gestures using the individual MFs, and outputted the average of 99.2% accuracy. The results show that the system achieved a high recognition performance by using the individual MFs.

5.2.2 Gesture recognition using the other signers membership functions (MFs)

In the second experiment, the gestures given by the signers were recognized using other signer's MFs. The experiment was done to evaluate the effect of individuality factors to the performance of the recognition system. Figure 12 shows the recognition results of the gestures given by a signer-1 and signer-4 by using a signer-5's MFs. Table 2 shows the results of the recognition of all 5 signers using other signer's MFs, and the average of the accuracy decreased to 84.0% compared with the result using the individual MFs, which was 99.2%. The results show that the accuracy of the system affected by individuality factors of the signers. On the other hand, gestures vary between individuals, and signers need a prior training for achieving high accuracy recognition.

5.2.3 Gesture recognition using universal membership functions (UMFs)

In this experiment, the gestures given by the signers were recognized using UMFs and Fig. 13 shows the results of the recognition. In the graph, five gestures were recognized with the accuracy less then 80%, which were "America", "Torture", "Meeting", "Same" and "Round" due to the complexity of the gestures. Moreover, the results show the gestures were composed of similar segment of movements and velocities. To compare the performance of the recognition system, an evaluation was examined by averaging the output accuracies obtained by using different type of the MFs for recognizing gestures as shown in Fig. 14. The graph shows the average of the accuracy increased to 94.1% compared with the accuracy using other signers MFs, which was 84.0%. Figure 15 shows the recognition results of a signer-2 and signer-3 using the membership functions generated by different signers. By using the UMFs, the accuracy of the system was improved compared with the results using other signer's MFs. The results show the robustness of the UMFs for the recognition of the gestures of various signers.

In this paper, the experimental results show that the proposed approach is capable of recognizing various gesture patterns. The main purpose of the study was to prove the ability that the proposed method could recognize gestures with a guaranteed

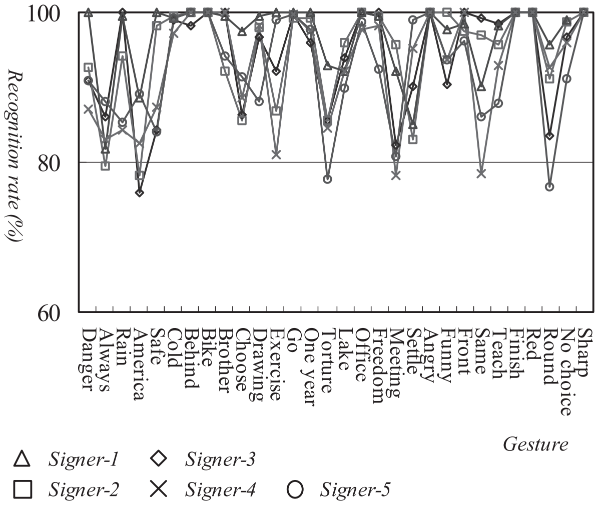


Fig. 13 Recognition results of all signers using UMFs.

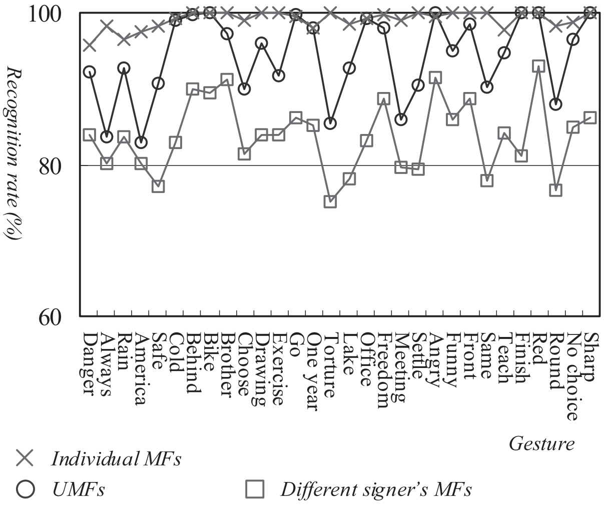


Fig. 14 The average of the recognition results.

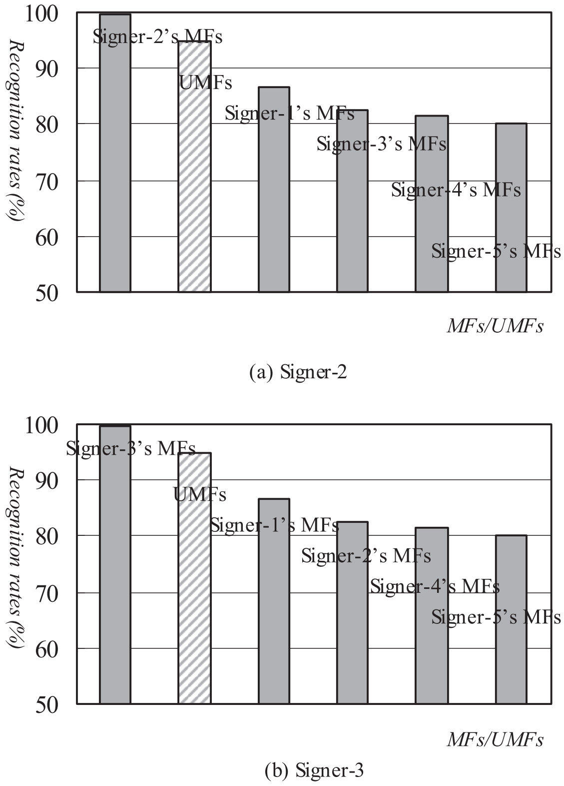


Fig. 15 Recognition results of the signer-2 and signer-3 using the various membership functions (MFs and UMFs).

accuracy. In order to increase the accuracy of the recognition system, a further investigation would be needed to reduce the effects of the individuality factors such as a body structure cue and the changing intentions in arm movements. Based on the size of the body structure, an adaptive recognition system could be developed for the purpose of choosing suitable MFs to be used as a template for recognizing gestures.

6. Discussion

In the human-computer interaction, it is important for machines to be able to communicate with humans not only using verbal media, but also non-verbal media for understanding them through body movements and facial expressions. The use of gestures would be one of the solutions in building an efficient HCI and its applications ranging from sign language recognition through medical rehabilitation to virtual reality. In this study, we are trying to analyze human characteristic body motions, and to recognize gestures to be used in HCI for the improvement of the communication between humans and machines. Our system employs a fuzzy probability technique for classifying features from the motions of human arms without considering shapes of the hands. The proposed gesture recognition algorithm is capable of reducing the affect of emotional factors and unconscious motions of a signer by employing the resampling algorithm for the normalization of gesture patterns. Additionally, motion features obtained from arm trajectories will work effectively for distinguishing hand gestures that are similar in the shapes and locations of hands [5]. Therefore, the development of a hybrid gesture recognition algorithm to fuse the information of the hand shapes, hand locations, palm orientations and arm motions is a potential study in order to increase the accuracy of the recognition systems.

Beyond the accuracy of the gesture recognition algorithm, the robustness of the recognition should be important issue to the adequate measurement in various environments. Recognizing human gestures is a challenging work due to the dynamic movements of a human body. Additionally, gestures vary among individuals, and are affected by emotional factors of signers. A large number of physical quantities are required to describe the physical components of human gestures. Hence, the use of quantitative representation is not enough in building a robust gesture recognition system. In other words, qualitative representation is needed for mapping physical quantities of human gestures as a symbolic function in order to recognize gestures in various environments. A qualitative representation is a knowledge description that could be employed to construct higher-level cognitive functions for the recognition system to reason and react in unpredictable environments, and it should be incorporated into system performance [20]. It means that the relation of quantitative and qualitative representations will be important for the design of the recognition system that performs intelligently without reducing its accuracy. The use of artificial intelligence (AI) for representing the qualitative reasoning is a potential researching direction for the development of a gesture recognition system that could be possibly used to practical applications.

7. Conclusion

In this paper, we introduced a gesture recognition approach by using a vision-based motion capture system. Our system employed a fuzzy probability technique for classifying features from the motions of human arms. The experimental results showed that the system had a good performance to recognize 30 sign languages by various subjects. The interesting point according to the proposed system was a gesture recognition, which was based on the arm trajectory without considering shapes of the hands. For recognizing the gesture, especially in the sign language application, the extraction of hand shapes might be important. However, based on the results of the experiments, features from arm motions could also be sufficiently used. Thus, the capability of the increase of other features would improve the recognition ability, and would give a significant innovation to the vision-based motion capture system, which provides an attractive communication tool for the interaction between humans and computers. In addition, the proposed system would work useful for the interaction between disabled people and computers, especially for a person that had a damage on the hand. The authors are planning to construct a system to understand gestural commands for supporting disabled people, which will be tested in the next stage.

References

[1] C. Manresa, J. Varona, R. Mas, and F.J. Perales: Hand tracking and gesture recognition for human-computer interaction, *Electronic Letters on Computer Vision and Image Analysis*, Vol. 5, No. 3, pp. 96–104, 2005.

[2] L. Brethes, P. Menezes, F. Lerasle, and J. Hayet: Face tracking and hand gesture recognition for human-robot interaction, *Proceedings ICRA*, Vol. 2, pp. 1901–1906, 2004.

[3] C. Manresa, J. Varona, R. Mas, and F.J. Perales: Hand tracking and gesture recognition for human-computer interaction, *Electronic Letters on Computer Vision and Image Analysis*, Vol. 5, No. 3, pp. 96–104, 2005.

[4] H.G. Okuno, K. Nakadai, and H. Kitano: Social interaction of humanoid robot based on audio-visual tracking, *Proceeding of the 18th Intern. Conf. Industrial and Engineering Applications of Artificial Intelligence and Expert Systems, IEA/AIE*, pp. 140–173, 2002.

[5] M.C. Su, Y.X. Zhao, H. Huang, and H.F. Chen: A fuzzy rule-based approach to recognizing 3-D arm movements, *IEEE Transaction on Neural System and Rehabilitation Engineering*, Vol. 9, No. 2, pp. 191–201, 2001.

[6] B. Min, H. Yoon, J. Soh, Y. Yang, and T. Ejima: Hand gesture recognition using hidden Markov models, *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 5, pp. 4232–4235, 1997.

[7] M. Aoba and Y. Takefuji: Motion feature extraction using second-order neural network and self-organizing map for gesture recognition, *Information Processing Society of Japan (IPSJ)*, Vol. 2004, No. 130 (20041220), pp. 5–8, 2004.

[8] G.S. Schmidt and D.H. House: Model-based motion filtering for improving arm gesture recognition performance, *Gesture-Based Communication in Human-Computer Interaction*, Lecture Notes in Computer Science, Vol. 2915/2004, pp. 210–230, 2004.

[9] S.S. Fels and G.E. Hinton: Glove-talk II-A neural-network interface which maps gestures to parallel format speech synthesizer control, *IEEE Transactions on Neural Network*, Vol. 8, No. 5, pp. 977–984, 2005.

[10] J. Lementec and P. Bajcsy: Recognition of arm gestures using multiple orientation sensor, *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*, pp. 965–970, 2004.

[11] H. Hienz, K. Grobel, and G. Offner: Real-time hand-arm motion analysis using a single video camera, *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pp. 323–327, 1996.

[12] A.F. Bobick and A.D. Wilson: A state-based approach to the representation and recognition of gestures, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 12, pp. 1325–1377, 1997.

[13] J. Yamato, J. Ohya, and K. Ishii: Recognizing human action in time-sequential images using hidden Markov model, *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognition*, pp. 379–385, 1992.

[14] J. Davis and M. Shah: Visual gesture recognition, *Visual Image Signal Processing*, Vol. 141, pp. 101–106, 1994.

[15] F. Bevilacqua, F. Guédy, N. Schnell, F. Emmanuel, and N. Leroy: Wireless sensor interface and gesture-follower for music pedagogy, *Proceedings of the Conference on New Interfaces for Musical Expression (NIME07)*, pp. 124–129, 2007.

[16] T. Starner, J. Weaver, and A. Pentland: Real time american sign language recognition using desk and wearable computer based video, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, pp. 1371–1375, 1998.

[17] H. Sawada, T. Ukegawa, and E. Benoit: Robust gesture recognition by possibilistic approach based on data resampling, *International Workshop on Fuzzy Systems & Innovational Computing (FIC2004)*, pp. 168–173, 2004.

[18] K. Wan and H. Sawada: 3D Motion prediction of human upper body by tracking reflective markers on a moving body, *Malaysia Japan International Symposium on Advanced Technology (MJISAT2007)*, MJISAT-172, 2007.

[19] L.A. Zadeh: Fuzzy Sets, *Information and Control*, Vol. 8, pp. 338–353, 1965.

[20] H. Liu: A fuzzy qualitative framework for connecting robot qualitative and quantitative representation, *IEEE Transactions on Fuzzy Systems*, Vol. 16, No. 6, pp. 1522–1530, 2008.

Khairunizam WAN

He received his B.E. degree in Electrical and Electronic Engineering from Yamaguchi University in 1999. He is currently undergoing Ph.D. course study at the Department of Intelligent Mechanical Engineering System, Faculty of Engineering, Kagawa University, Japan since April 2006. He is a member of IEEE.

Hideyuki SAWADA (Member)

He is an Associate Professor in the Faculty of Engineering, Kagawa University. He received the B.E., M.E. and Dr. Eng. degrees in Applied Physics from Waseda University in 1990, 1992 and 1999, respectively. He was a Research Fellow of the Japan Society for the Promotion of Science from 1995 to 1998. His current research interests include sound and image processing, neural networks and robotics. He is a member of IEEE, IPSJ, IEICE, RSJ, and JSME.