K-18

# Avatar Language: Enhancing Natural Human Communication and Intercultural Understanding

*Jin Hou and Yoshinao Aoki*

## 1. Introduction

There has been a growing interest in building natural human communication system. Nonverbal language plays an important role in natural human interaction, because it can accompany verbal communication, allowing the avatar to strengthen, make explicit or weaken the sense of their expression, and to enable the interlocutors to give each other information about their identity or their environment. Especially, it is a promising way to help understand across the intercultural barrier. So the implementation of Human-Computer Interface allowing users to employ nonverbal language during their interactions is important for the quality of the interactions. This paper proposes to design this kind of nonverbal language for 3-D avatars, so called avatar language [1], and implement a novel real-time avatar language communication system from natural language instruction by an intelligent communication method. Experiment performed via the Internet show this avatar language chatting system is a promising natural human communication system, and a potential way of bridging the gap between different cultures or countries.

## 2. Parameterized Action Representation Based on the Anatomical Structure of a 3-D Model

### 2.1. Static Parameterized Representation of the Shape of Hands and Arms

According to the anatomical structure, human hand skeleton consists of 27 bones. We classified the hand joints into three types: flexion or twist, directive and spherical joints, which have 1 degree of freedom (DOF) (extension/flexion), 2 DOFs (one for extension/flexion and one for adduction/abduction) and 3 DOFs (rotation) respectively. Based on this theory, each finger (from index to pinky) has four DOFs, while the thumb has five DOFs. We describe these DOFs by the joint angles $\phi_{ij}$. In addition to the 6 DOFs for the position and rotation of the wrist, the hand model has 27 DOFs in total. Therefore, the shape of hand could be decided by the joint angles of fingers. We describe the hand shape (HS) by the following formulas.

$$HS = F(\Phi) = F(\Phi_0, \Phi_1, \Phi_2, \Phi_3, \Phi_4) \qquad (1)$$

$$\Phi_i = \begin{cases} (\phi_{i0}, \phi_{i1}, \phi_{i2}, \phi_{i3}, \phi_{i4})^T & if \;\; i=0 \\ (0, \phi_{i1}, \phi_{i2}, \phi_{i3}, \phi_{i4})^T & if \;\; i=1\sim4 \end{cases} \qquad (2)$$

While the arm shape (AS) representation could be described by the same method:

$$AS = F(\Theta) = F(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9) \qquad (3)$$

Department of Electronics and Information Engineering,
Graduate School of Engineering, Hokkaido University

### 2.2. Dynamic Parameterized Representation of the Movement of Hands and Arms

In this paper, we propose a complex algorithm of vector rotation and addition. We explain this method by giving an example of a finger.

Firstly, we specify a vector for each segment of a finger. We define local coordinate systems of the finger on every joint position with conventional right-handed coordinate systems. Each joint's rotation R is expressed by a sequence of rotations occurring around the x, y, z axes of the local coordinate system R=Rx (α) Ry (β) Rz (γ), where α, β, γ are joint rotation angles around x, y, z respectively. Secondly, we calculate the joint motion by rotating the segment vector at the proximal joint and then adding it to the parent segment vector to get the expression in the global coordinate. This calculation is shown in Eq. (4).

$$\begin{cases} \vec{E}_1 = R_1 \vec{E}_{01} \\ \vec{E}_2 = R_1(\vec{E}_{01} + R_2 \vec{E}_{02}) = \vec{E}_1 + R_1 R_2 \vec{E}_{02} \\ \vec{E}_3 = R_1\left[\vec{E}_{01} + R_2\left(\vec{E}_{02} + R_3 \vec{E}_{03}\right)\right] = \vec{E}_2 + R_1 R_2 R_3 \vec{E}_{03} \\ R_i = R_{ix}(\alpha) R_{iy}(\beta) R_{iz}(\gamma) \qquad i=1\sim3 \end{cases} \qquad (4)$$

This complex algorithm is suitable for the movement of arms as well, in that case, the upper arm vector and forearm vector are used. Hence, we can describe the movements of hands and arms by the coordinate values of a sequence of vectors.

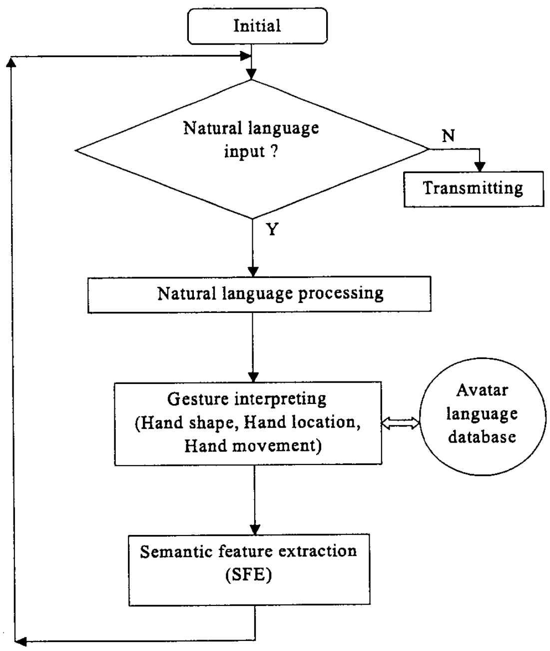


Fig.1 The flow chart of the semantic feature extraction

Fig.2 Communication with realistic sensations in avatar language

### 2.3. Semantic Feature Detection From the Hand Shape, Hand Location and Hand Movement

Referring to sign languages [2][3], for avatar languages, we classify 21 kinds of hand shapes (coded as HS01~HS21), and define 30 locations (coded as HL01~HL30). A visual avatar language is expressed by a sequence of actions, i.e., animations of avatars. Animations are described by the start location, the end location and the moving path. The moving path can be worked out by some interpolators between the start point and the end point. In other words, if we know the start and the end locations of a hand, we can interpolate some points to the route according to the movement of a hand, so that animations can be described very precisely. We define 18 kinds of movement (coded as HM01~HM18) for a hand such as rotation, shaking, bending. Therefore, while we build an avatar language database based on this notation where some parameters could be calculated offline in advance, the parameters (semantic feature) for avatar language real-time description could be indexed rapidly online. Fig.1 shows the flow chart of the semantic feature extraction.

## 3. Avatar Language Communication Experiment in Networked Virtual Environment

Our purpose is to build a real-time avatar language chatting system in networked virtual environment where the avatars who represent the human beings can communicate with each other in avatar language in the real time, with so much sense of presence that the user may have the illusion of being in the real world [4].

The client-server architecture is constructed in Java to make the avatars share the same VRML world. The major roles of the server are receiving the information from the client and sending it to the other clients, while the responsibilities of the client are sending the information of it to the dispatcher and receiving the information about the other clients from the dispatcher. Each client that receives the same VRML file is connected to the same server so that the shared cyberspace is generated. The user will transfer the event to the server through the browsers if any change happens, and the server will notify that to the other users in the shared space immediately so the animations will be shared by all users.

The scenario of communication between the server and the client is as follows: (1) Start a server. (2) The server is waiting for a request from the clients. (3) A client reads a VRML file through the browser and is connected to the server automatically. (4) The user in this client animates or moves the avatar representing himself and sends this event to the server through the browser. (5) Another client loads the same VRML file and is connected to the same server automatically too. (6) The clients exchange their information with each other through the server and share the networked virtual environment.

An example of experiments is shown in Fig. 2, where the Japanese animation means "I want to learn avatar language", while the Chinese animation means "I teach you".

## 4. Conclusion

This paper proposes a new conception, avatar language, which is a visual nonverbal language for natural human interaction and intercultural understanding. Communication system experiments using avatar language are implemented via the Internet. Different from other current gestural chatting systems, our system gives avatars run-time instructions from natural language. Remote participants can contact each other with both the text and animated embodied agents in real-time, which makes the communication more interesting, explicit and natural. For efficient delivery, we use a novel natural language processing and intelligent communication method: extract the semantic information data that is a sequence of parameters from the input text, and transmit the limited data only for animating the 3-D hierarchical avatars.

## 5. Acknowledgement

This work is partially supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan under Grant 13305026.

## 6. References

[1] J. Hou and Y. Aoki, "Design and experimentation of avatar language," *Proc. IEEE Sapporo Section Student Paper Contest*, pp.420-421, Hokkaido Tokai University, Sapporo, Japan, Oct.2001.

[2] A. Yonekawa, *Research on Sign Language Description*. Meiji Study, Japan, 1984.

[3] Chinese Deaf Association, *Chinese Sign Language*. Huaxia Publishing, China, 1990.

[4] R. Lea, K. Matsuda and K. Miyashita, *Java for 3D and VRML Worlds*. Prentice Hall, Japan, 1997.