Real-time Human Proxy : An Avatar-based Interaction System

Arita, Daisaku
Department of Intelligent Systems, Kyushu University

Yoshimatsu, Hisato
Department of Intelligent Systems, Kyushu University

Hayama, Daisuke
Department of Intelligent Systems, Kyushu University

Taniguchi, Rin-ichiro
Department of Intelligent Systems, Kyushu University

https://hdl.handle.net/2324/5868

出版情報：Proceedings of Second International Workshop, IMTCI 2004, pp.13-18, 2004-09
バージョン：
権利関係：

Real-time Human Proxy: An Avatar-based Interaction System

Daisaku Arita, Hisato Yoshimatsu, Daisuke Hayama, and Rin-ichiro Taniguchi

Kyushu University, Kasuga, Fukuoka 816-8580 Japan,
{arita,hisato,daizz,rin}@limu.is.kyushu-u.ac.jp,
http://limu.is.kyushu-u.ac.jp/

1 Background

There are several researches on virtual environments for human interaction. In these researches, a 3-D virtual space is reconstructed, in which each participant is represented as an avatar by computer graphics techniques. Through the reconstructed virtual space, each participant sees and hears other participants' activities from the position where his/her avatar is represented, where their positional relations are consistent virtually. This means that each participant can understand where other participants stand, look at, and point to, or understand where a sound comes from, and also move around in the virtual space. However, these avatar-based interaction systems have difficulty in controlling avatars, where the degrees of freedom of human body are so high that legacy input devices are not sufficient to acquire or input participants' activities. In this paper, we describe Real-time Human Proxy; our concept to provide realistic virtual-space-based communication.

2 What is Real-time Human Proxy?

In avatar-based interaction, an avatar is expected to reflect activities of a participant into a virtual space as if he/she were there. In order to achieve this requirment, we have developed a vision-based motion capture system (MCS)[1] as an input device without compelling participants annoying operations.

Although using an MCS as an input device, it is not possible to reflect all of participant's motions into an avatar since an MCS cannot acquire all information such as shape of hands, or movement of eyebrows and a lip. Just feeding captured data into an avatar, the lack of information may cause unnatural avatar motions. On the other hand, it is not necessary for an avatar to act exactly the same as participant's motions since participants usually do not want to know how others moving but what others are doing.

Real-time Human Proxy (RHP) is a new concept for avatar-based interaction, which virtualizes a human in the real world in real-time. It makes better use of an MCS and makes avatar act more meaningfully. The aim is to make an avatar act as if he/she in a distant place is present in a virtual space. Therefore, we focus on acquisition and representation of human action or nonverbal information. We

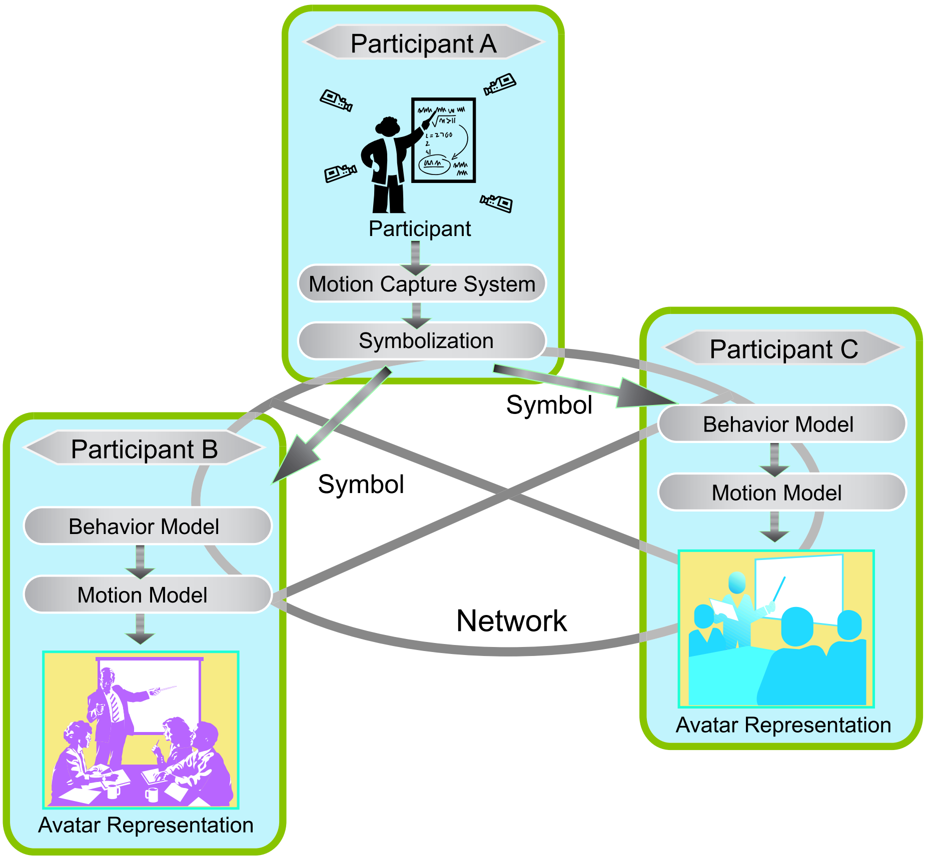


Fig. 1. The concept of RHP

use symbolization on acquisition, and use *Motion Model* and *Behavior Model* on representation, that they are illustrated in Fig.1 and described below.

On RHP, we acquire human actions instead of human motions. We segment and categorize motion sequences into pre-defined actions, expressing them as symbols. Each symbol is formed by a label of an action and its parameters, such as "walking (ν_x, ν_y)" where (ν_x, ν_y) is the velocity of participant. After recognizing human actions from captured motion data, the system transmits the symbols to the representation side of the virtual space.

Since each symbol transmitted from a participant is formed as a label of an action, to make an avatar act, pre-defined knowledge is needed, which is a set of trajectories or motion sequences of each body part, tied to each symbol. We call the knowledge *Motion Model*.

We have another model called *Behavior Model* for describing the human actions (see Fig.2). Behavior model is a set of state transition graphs that decides

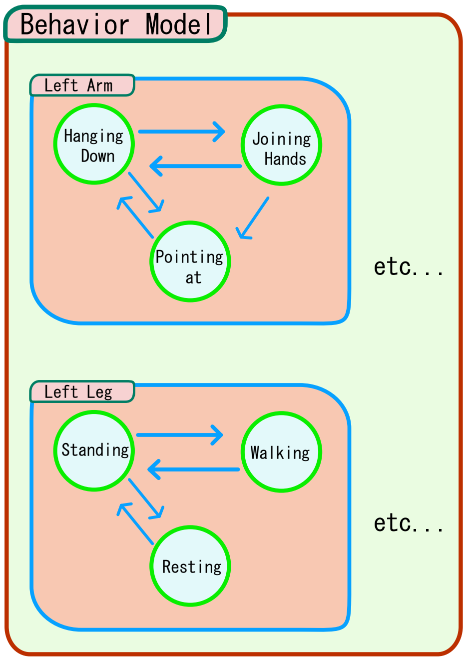


Fig. 2. Example of Behavior Model

the next action an avatar is going to do. Each graph corresponds to a body part such as a right arm and left leg, and each state in the graphs corresponds to an action or a symbol such as "walking", "raising hand", and "pointing with finger." When a symbol is transmitted, the current state of the graph is forced to transit to the state corresponding to the symbol. In addition, an avatar often freezes if the avatar acts only when symbols are transmitted, since no symbols are transmitted when a participant does not make any pre-defined actions. Needless to say, such avatar's behavior does not seem natural. To solve this problem, Behavior Model has some actions invoked spontaneously such as "folding arms" or "sticking hand into a pocket." These actions may work for time filling, and may make participants feel more natural[2]. And of course, these actions do not indicate participant's intentions in order not to influence interaction between participants.

Behavior Model can maintain consistency in transition of action. Unnatural transition of actions will not be provided, where any transition from an action to another is a production of transition probabilities in Behavior Model.

Using the above models, an avatar behaves in a virtual space according to transmitted symbols. At first, Behavior Model changes transition probabilities in its transition graphs so that the avatar's state transits to the state corresponding to the symbol just received immediately. Secondly, the avatar performs an

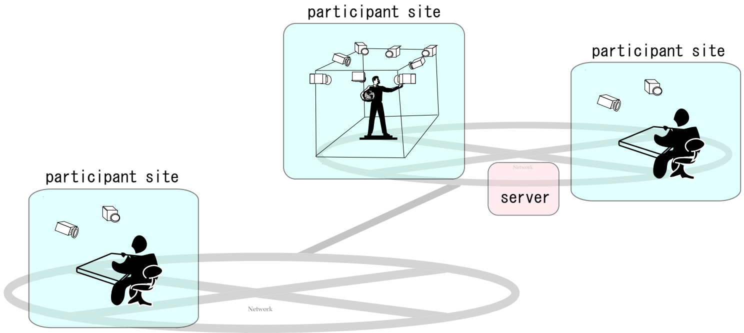


Fig. 3. System configuration of the game: Three participants located separately are connected by LAN.

action referring to the Motion Model tied to the state just have been activated in Behavior Model. A participant is able to see a virtual space in which any participants, including him/herself, are represented as avatars. from the viewpoint his/her avatar.

3 Benefits of RHP

The concept, i.e., symbolization of actions, and representation of an avatar using Motion Model and Behavior Model, makes representation process simple. As described above, acquired information about participant's motion does not have detailed information of several body parts such as angles of fingers, which must be compensated to provide natural representation. Using RHP, actions are already recognized, and it is possible to generate full body motion sequences in an arbitrary manner referring to Motion Model.

Moreover, this allows avatars to be designed beyond constraints of avatar's physical structure. Ordinarily, each avatar is designed to fit the participant, so that 3-dimensional positions of body parts acquired from an MCS is correctly represented in computer graphics. This requires that the sizes of body parts are strictly same between an participant and its avatar. On the other hand, RHP transmits the symbolized action, not the 3-dimensional positions directly. Consequently RHP largely relaxes the constraints. We can use any kind of avatar including higher body, shorter arms, bigger head. Moreover, an avatar does not need to be a human. It may be a fish for example. This improves not only usability of the system, but also variety of avatars in interaction.

4 Experimental Results

We have developed a simple interactive game system with RHP to verify whether participants can naturally interact with each other in case that all actions necessary for interaction can be listed to recognize and display. There are three

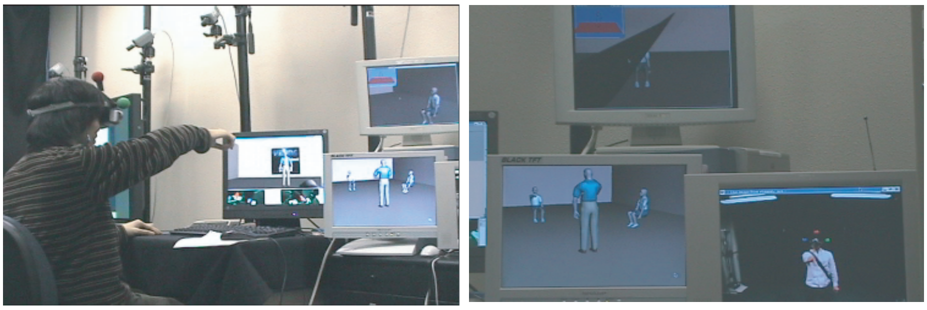


Fig. 4. A scene of interaction: The left image shows a participant with an HMD and two cameras. Motions of his head and arms are captured. The right bottom display in the right image shows another participant with an HMD and surrounding cameras. Motions of his whole body are captured. The upper display shows view of the participant. The left bottom display shows the overview of the virtual space.

participants located separately as shown in Fig.3. Motion information about each participant is captured by a full-body motion capture system[1] or a upper-body motion capture system[3] we have developed. Each participant wears a HMD to see what he/she want to see in a virtual environment.

The rules of the game, which is a simplified version of a famous game in Japan, are as shown in Fig.5 with some snapshots. Symbols transmitted among participants are “finger pointing”, “hands up”, and “head turn.” In addition, voice information is transmitted via microphones and headphones.

The impressions of participants are as follows:

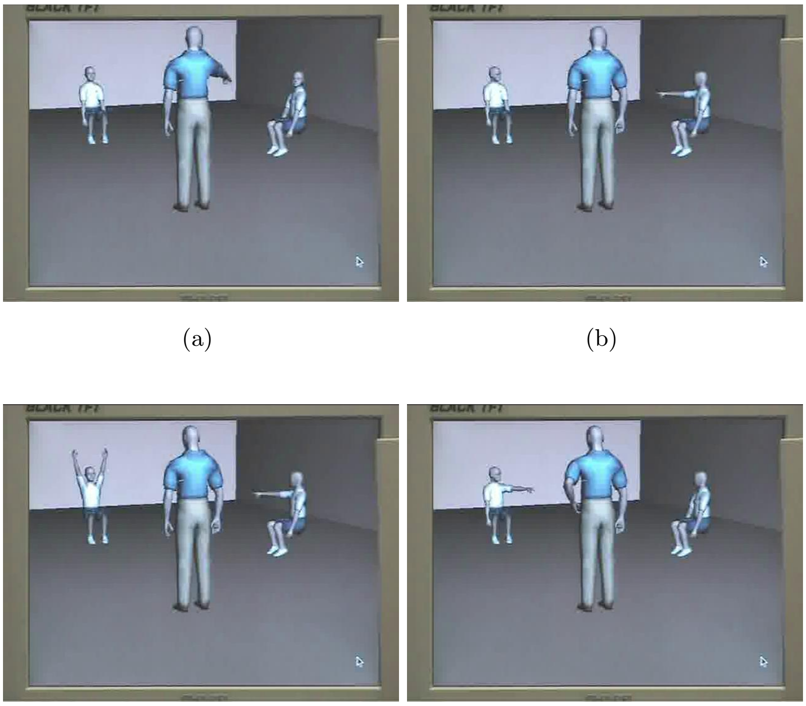
- At the beginning of interaction, participants doubted that avatars were controlled by other participants since they was able to see only avatars represented in computer graphics. However, such feeling became fainter in a short time since avatars acted in time to the game flow.
- Each participant felt that others looked at him/her since he/she saw that avatars turned their heads.
- Participants were able to easily understand where avatars looked and pointed.

5 Conclusion

In this paper, we propose a concept of Real-time Human Proxy for avatar-based interaction systems, which virtualizes a human in the real world in real-time, and which makes the virtualized human behave as if he/she were present. To verify the effectiveness of RHP, we have developed an interactive game system.

References

1. Date, N., Yoshimoto, H., Arita, D., Yonemoto, S., Taniguchi, R.: Performance evaluation of vision-based real-time motion capture. In: Proc. of Workshop on



(a) (b)

(c) (d)

Fig. 5. An interactive game: One of participants becomes a leader. (a)The leader says "A" and points to one of participants. (b)The participant who is pointed to at the previous step says "B" and points to one of participants. (c)The participant who is pointed to at the previous step says "C" and puts his/her hands up. (d)The participant who puts his/her hands up becomes a leader of the next turn until someone fails.

Parallel and Distributed Computing in Image Processing, Video Processing, and Multimedia, in IPDPS CD-Rom Proceedings. (2003)

2. Arita, D., Taniguchi, R.: Non-verbal human communication using avatars in a virtual space. In: Proc. of International Conference on Knowledge-Based Intelligent Information and Engineering Systems. (2003) 1077–1084
3. Yonemoto, S., Taniguchi, R.: A direct manipulation interface with vision-basedhuman figure control. In: Proc. of HCI International. (2003) 811–815