Capture and Express Behavior Environment (CEBE) for Realizing Enculturating Human-Agent Interaction

Yoshimasa Ohmoto, Akihiro Takahashi, Hiroki Ohashi, and Toyoaki Nishida

Graduate school of Informatics, Kyoto University
Yoshidahonmachi, Sakyo-ku, Kyoto-shi, Kyoto-hu, Japan
{ohmoto@,a.takahashi@ii.ist,ohashi@ii.ist,nishida@}i.kyoto-u.ac.jp

Abstract. We are studying how Embodied Conversational Agents (ECAs) express communication behavior with cultural background. The objective of this study is the proposition of the modified Capture and Express Behavior Environment (CEBE) in which a person can interact with ECAs controlled by the captured behavior of another person with cultural background. In this paper, we discuss modifications and concepts of CEBE to apply CEBE for investigations to realize an ECA with cultural background. The prototype system could capture basic human behavior, such as head direction, posture of the upper body, and 3D angles of arms, when each part of the body, such as head, hands, arms and trunk. In addition, the system could control a robot or a virtual agent based on the detected data. We have to develop some implementations to interact with people with cultural background.

Keywords: human-agent interaction, measuring system, WOZ system, Robot operation interface.

1 Introduction

In recent years, there has been a growing exchange in goods, information and people in the world amidst the progression of internationalization. In cross-cultural exchange, people can relativize their own values and obtain new values from different cultures. Although there are some cross-cultural differences in habits and rules, the experience of recognizing these differences gradually is important in acquiring communication skill. The reason why people can get though them is that they have great abilities to learn, adapt and communicate with other people.

On the other hand, Embodied Conversational Agents (ECAs) which are expected to provide an effective interface for many users, including computer novices in diverse fields ([1], [9], [12]), however, have such ability in the distant future. It is important to interact with people in the fields. When ECAs interact with people, the cross-cultural features influence human-agent interactions. ECAs can have databases of cross-cultural rules and habits which are known well in advance. There are however body motions, behavior, and mental stances

T. Ishida (Ed.): Culture and Computing, LNCS 6259, pp. 41–54, 2010.
© Springer-Verlag Berlin Heidelberg 2010

which influence cross-cultural interactions, and it is hard to describe them as rules or habits. These differences appear not only in languages and their use, but also in facial expressions, face directions, gestures, the range of movements, postures, interpersonal distance, and so on [8].

We are studying how ECAs express communication behavior which is not known as rules in a culture, such as emotional representation, the degree of agreement or disagreement, and commitment to the communication, based on cross-cultural background. In this paper, "cultural background" means meaningful customs, habits, gestures, nonverbal information and mental stances in a particular culture. We expect that the cultural background influences communication behavior. For the studies, we pay attention to gaze (head orientation), postures, hand gestures and paralinguistic features (pitch and power of voice). We will investigate them in a negotiation situation in which people have to express their opinions and speculate the inner states of their communication partner.

Cultural behavior is observed in diverse situations and representations. Different layers of phenomena exist that are influenced by cultural aspects such as the verbal and nonverbal behavior, appearance, proxemics, learning strategies and many more. We thus have to consider different appearance and physical limitations when ECAs express cultural behavior. Moreover, it is hard to pinpoint down the constituents of culture and its effect on interaction [13]. Therefore, we have to design ECAs based on data which is obtained in actual human-agent interaction. It is hard to apply the foundations of cross-cultural communication by analyzing data obtained in human-human communication to ECAs which have different appearance and physical limitations than humans.

We, however, did not have an environment to obtain the data in actual human-agent interaction. In other words, we need to analyze enculturated agent-human interaction in an environment in which a human can interact with another human who has the appearance and physical limitations of ECAs.

The purpose of this study is to propose an environment in which a person can interact with ECAs controlled by captured behavior of another person with cultural background, so that we can analyze the data obtained in actual enculturated agent-human interaction. The proposed environment is called Capture and Express Behavior Environment (CEBE). Human behavior in CEBE is expressed by an ECA which has a different appearance and physical limitations from a human. If the ECA could express enculturated behavior by CEBE, a standalone ECA which has the same appearance and physical limitations could express enculturated behavior without pinpointing down the constituents. Therefore, we expect that CEBE is a useful tool to analyze cross-cultural behavior expressed by ECAs. We discussed the prototype of CEBE in general situations in another paper [10]. In this paper, we introduce the basic architecture and the prototype, and we discuss modifications and concepts of CEBE to apply CEBE for investigations to realize an ECA with cultural background.

CEBE will be used as a Wizard of OZ (WOZ) system in two types of research. The first is data capturing to realize agents which have a communication skill with existing cultural background for cross-cultural communication. The other is the investigations of human-agent interaction on the supposition that humans and agents which controlled a human belong to different culture.

The rest of the paper is organized as follows. Section 2 reviews related works and discusses achievements and non achievements of previous work. Section 3 explains the design and architecture of CEBE. Section 4 describes implementations to detect human behavior with cultural background in CEBE. Section 5 discusses concepts and limitations of CEBE to capture and analyze human behavior with cultural background. Finally, Section 6 contains conclusions and future works.

2 Previous Works

A number of research groups have studied the use of ECAs for intercultural communication.

Huang et al. [6] developed a culture-adaptive virtual tour guide agent that is implemented in a modular way with the GECA Framework [5] to minimize the development cost. It can switch its behavior and speech language to three culture modes: general Western, Japanese, or Croatian. This study focuses only on the surface traits of culture, that is, languages, symbolic gestures, and probably culture-dependent characteristics of gestures. Since they used scripts to describe human-agent interactions, the range of possible interactions will be relatively limited and the quality of the whole system heavily depends on the knowledge and skill of the agent designers.

Iacobelli and Cassell [7] showed an attempt to use virtual peers to encourage African American children to switch their language coding to increase school-based literacy. They implemented two virtual peers by applying two models of behavior to an existing virtual child with a racially-ambiguous appearance in their previous study [2]. They implemented their enculturated behavior on the basis of observations of the verbal and non-verbal behavior of children. Their agents, therefore, had the same limitations as an agent in Huang et al [6].

Some researchers developed their systems to generate culturally adequate behavior based on a human's observable behavior instead of implementations of cultural backgrounds by heuristics. For example, Rehm et al. [14] proposed a Bayesian network model of cultural adaptation, which was then employed in two different sample applications that illustrate the great potential of culture adaptive systems. They, however, realized the model with empirical data from the corpus study under standardized conditions at hand. In addition, the empirical data was cross-cultural rules and habits which could be known in advance. This is a common issue when a behavior model is developed from empirical data at hand.

There is another method to learning cross-cultural behavior, which is through by behavior data obtained from encultured human-agent interaction in which an agent directly expresses behavior of a human operator. This method is our approach. There are some systems in which an agent directly reflects human actions.

Ghobadi et al. [3] proposed a system to control a hand based robot using 2D/3D images. These images are used as the input for hand detection, classification and a tracking system which is used as an interface for sending the commands to an industrial robot. Hirsch et al. [4] also developed BiDi (bidirectional) screen to support 2D multi-touch and walk-up 3D gesture interaction. An image sensor placed a small distance behind an LCD forms a mask-based light-field camera, allowing passive depth estimation. The estimated scene depth can be used to support real-time 3D gesture interaction. These systems are specialized for hand operations, and they could not capture cross-cultural behavior in human-agent interaction.

Tachi et al. [15] had developed a master-slave manipulation system with the function of mutual telexistence (TELEsarPHONE). The system composed of 3 subsystems; a slave robot TELESAR II, a master cockpit, and a viewer system. The operator sitting in the cockpit manipulated the slave robot through two master arms, two master hands multi-stereo display system, speakers and a microphone. The operator had to equip these master arms and master hands. By using this system, the robot reflects the arm and hand actions of the operator. The devices, however, would prevent natural interaction with a cultural background.

We have discussed achievements and non achievements of previous work for the purpose of our study. Some researchers presented ECAs for intercultural communication. Most ECAs were, however, implemented their cross-cultural behaviors and behavior models on the basis of observations of human-human interaction. Possible interactions of these ECAs are relatively limited and the quality depends on the agent designers. To solve this issue, there is a method to learning cross-cultural behavior by behavior data obtained from HAI, in which an agent reflects human behavior. Some systems where an agent reflects human actions could not achieve natural interactions with a cultural background.

The objective of this study is the proposition of CEBE design. This study is different from previous works in following ways; the agent reflects human actions without preventing natural interaction by using an immersive environment and non-contact sensors. In addition, a person in CEBE can intuitively understand the differences of appearance and physical capabilities between ECAs and human. These features are necessary to capture human behavior in natural interaction with a cultural background. In addition, we discuss modifications and concepts of CEBE to apply CEBE for investigations to realize an ECA with cultural background, and we partly implemented.

3 Framework

In this section, we propose a design and architecture of CEBE.

Two kinds of ECAs are expected to perform in CEBE; virtual agents and robots. Virtual agents are suitable for interactions in which expressive gestures are needed because most of them can express facial expressions, gaze directions and subtle movements. We had better use virtual agents when we focus on those features in cross-cultural communication. The virtual agents, however, are unsuitable for interactions in which ECAs have to interact with substantial objects in real world because most of them do not have their body in real world. On the other hand, interactions which are suitable for robots are vice versa. For example, we had better use robots to investigate a multiple-party interaction in which people in different cultures take part because virtual agents cannot join the multiple-party conversation. Both these kinds of ECAs are used depending on the situation. Therefore, we develop CEBE to be able to use both types of ECAs and we will investigate them in both situations.

3.1 Necessary Conditions for CEBE

CEBE is an environment in which a human can interact with another human who has the appearance and physical limitations of ECAs. CEBE has two aspects; one is a system to obtain data of cultural behavior in a human interaction through a body and perception of an ECA. Another is a system to operate a body of an ECA by using perception of the ECA and obtained data.

In either case, a person in CEBE has to be able to perceive the same level of information which an ECA can perceive. We need to consider the difference of information processing between ECAs and humans. For example, a human can recognize his/her surroundings by using only video in most cases, because visual information is easily interpretable by humans. ECAs, however, need lots of sensor data, such as a range sensor, 360-degree camera, microphone, and so on, to recognize their surroundings at the same level. It is important to provide both a person and an ECA with the same level of information. It is, however, difficult to process all of the information which is used in human-human communication. Therefore, we need to control the information provided to a person depending on the situation.

We also have to be able to obtain and express cultural interaction behavior in a human-agent interaction. For example, we behave in various ways, such as shaking our heads, making hand gestures and changing prosody, when we negotiate with others. We also pay attention to nonverbal behavior to speculate others' intentions. We expect that significant information which expresses cultural interaction behavior is gaze directions (head orientation), movements of the upper body and hands, and prosody of utterances (pitch and power of voices). The accuracy of measuring and reproducing behavior is dependent on the situation. We expect that ECAs cannot completely reproduce human behavior in most cases. We also expect that people can convert their cultural behavior which ECAs can reproduce by mutual adaptations, in the case that they understand the appearance and physical limitations of ECAs.

To sum up, necessary conditions for CEBE are as follows:

– Both a person and an ECA are provided with the same level of information which is needed to interact with cultural behavior.
– Cultural interaction behavior in a human-agent interaction can be obtained and expressed.

– A person in CEBE can intuitively understand appearance and physical limitations of ECAs, and he/she can convert their cultural behavior which ECAs can reproduce.

The accuracy of measuring and reproducing behavior is depending on the situation.

3.2 Prototype

On the basis of necessary conditions for CEBE, we developed a prototype CEBE which detects head direction, posture of the upper body, and 3D angles of arms, and controls a head and arms of a robot or a virtual agent based on the detected data. We currently implemented the methods to capture human behavior to apply CEBE for using robots. We thus did not implement the methods to capture gaze directions, facial expressions and hand gestures.

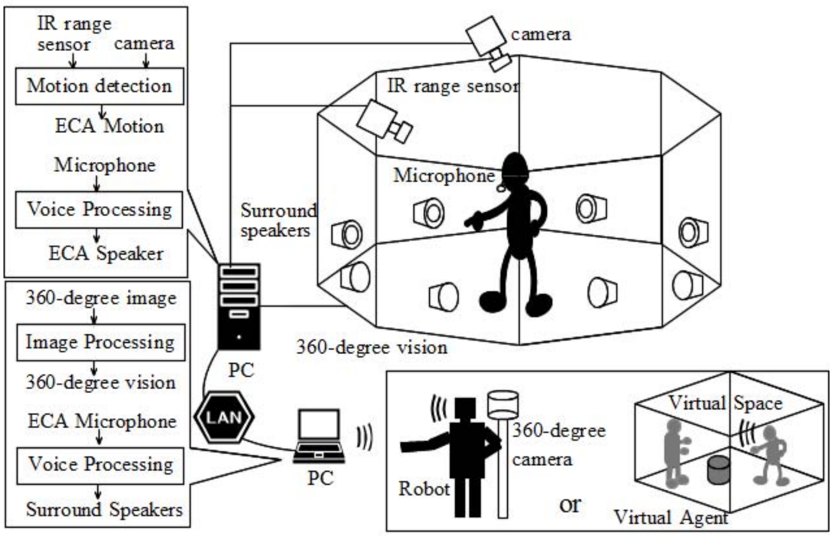

Fig. 1. An architecture of prototype CEBE

Devices which we used in the prototype. Figure 1 shows the whole image of the prototype system.

The prototype system had an immersive display, which reproduces 360-degree image around an ECA. The display intuitively provides an operator with surroundings of the ECA. We expected, therefore, that the operator's cognitive load to know the surroundings was reduced by using this display.

We used non-contact or small sensors to detect the operator's behavior in the immersive display because the sensors would not prevent natural interaction. Optical motion capture system is a typical non-contact sensor which can measure human body motion. The system needs many cameras for robust measuring. We, however, could not set enough cameras for robust measurement in the prototype system because the system used the immersive display in which there was small space. We therefore developed a motion capture system by using infrared range sensors (IR-range sensors) for the system.

ECAs which are used in this architecture are virtual agents which are implemented using the GECA Framework [5] and multi-degree of freedom robots (NAO, Aldebaran Co., Ltd). NAO is a robot which can be controlled by using wireless LAN and control programs. We can control angles of the head, shoulders, elbows, wrists, fingers, hip joints, knees and ankles of NAO.

(a) 360-degree camera (b) Immersive display (c) IR-range sensor (d) NAO

Fig. 2. Devices which we use in the prototype

Figure 2 shows the devices in the prototype system. The 360-degree camera and NAO is used when we use a robot as an operated ECA. The immersive display composed of eight portrait orientation LCD monitors with a 65 inch screen size in an octagon shape. The diameter of the octagon is about 2.5 meters. The infrared range sensor is used to detect human upper body motions.

System architecture. Figure 1 also shows an architecture of the prototype system which we are implementing. The prototype system can obtain and express head orientation, movements of upper body and hands, and prosody of utterances (pitch and power of voices) for the preliminary experiment, in which a negotiation task will be conducted.

An immersive display reproduces a 360-degree image around an ECA to provide a person in CEBE with information which the ECA perceives. When we use a virtual agent, the virtual space image which is around the virtual agent is displayed. When we use a robot, the image which is captured by the 360-degree camera is displayed. The person in the 360-degree vision system can see all around the ECA. Movements of the upper body and hands of the person are not prevented by the displays. Eight surround speakers play sounds which the ECA can listen. The person can detect the direction of the sound source to some degree.

Head orientation and movements of the upper body and hands of the person are captured by IR-range sensors (Swissranger, mesa imaging), acceleration sensors and image processing. When we use a virtual agent, a body of the virtual agent is moved by the captured data. When we use a robot, the robot is controlled by data which is converted to meet the limitations of robot's degree of freedom. The person can see the movements of the ECA in a window which is displayed in the 360-degree vision system. The voices of the person are captured by a headset microphone, and the ECA plays the voice. The voice passes through a voice processing component. The voice processing component can filter the voice depending on the purpose. When the voice is filtered, the person can hear the filtered voice.

Head orientation and movements of the upper body and hands of the person are captured by IR-range sensors (Swissranger, mesa imaging), acceleration sensors and image processing. When we use a virtual agent, a body of the virtual agent is moved by the captured data. When we use a robot, the robot is controlled by data which is converted to meet the limitations of robot's degree of freedom. The person can see the movements of the ECA in a window which is displayed in the 360-degree vision system. The voices of the person are captured by a headset microphone, and the ECA plays the voice. The voice passes through a voice processing component. The voice processing component can filter the voice depending on the purpose. When the voice is filtered, the person can hear the filtered voice.

4 The Method of Human Motion Capture

We discussed the prototype of CEBE in general situations in another paper [10]. In this section, we describe the modifications to capture human motion with cultural background in the prototype system. In this paper, we only explain the method to measure head rotation and angles of arms because upper body movements are important in the immersive environment. These implementations are currently in progress.

The prototype system could capture basic human behavior, such as a face direction and pointing, when each part of the body, such as head, hands, arms and trunk, were occluded below 30%.

4.1 The Implementation to Measure Head Rotation

In the CEBE for general situations, we implemented the method to measure head rotation by using an reflection intensity image (we call this as "amplitude") of IR-range sensors. This method was not suitable for measuring detailed degree of the head rotation angle. We implemented the method to measure head rotation by using a three-axis acceleration sensor and image processing because we expected that we had to measure detailed head motion for investigations about cultural differences. We could easily measure the vertical head rotation by using the three-axis acceleration sensor. In addition, the sensor would not prevent human

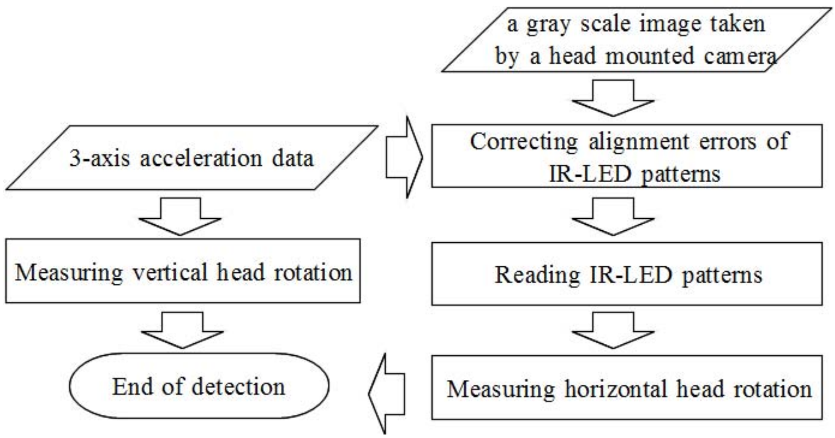

Fig. 3. The overview of the head rotation measuring process

actions because it was adequately small. Meanwhile, we measured the horizontal head rotation by using image processing, for which we used a head mounted camera and infrared-LEDs.

The overview of the measuring process is shown in Figure 3. The inputs of the measuring process are three-axis acceleration data and a gray scale image taken by a head mounted camera. The gray scale image captures the IR-LED patterns which are set between the LCD monitors composed the immersive display. The IR-LED patterns indicate the horizontal angle. In the first step, the angle of vertical head rotation is extracted from the three-axis acceleration data. In the next step, alignment errors of IR-LED patterns in the gray scale image are correcting by using the acceleration data. In the third step, the angle of horizontal head rotation is estimated by reading the IR-LED patterns.

Although the method to measure head rotation is robust, the head mounted camera made some people feel uneasy in test installations. We can measure head rotation by using a geomagnetic sensor, infrared range sensors and only image processing. We are developing another method to measure head rotation by using a small sensor which can measure geomagnetism and 3-axis acceleration because of reducing this uneasiness of users.

4.2 The Implementation to Estimate Angles of Arms

In the CEBE for general situations, we implemented the method to estimate angles of arms by using distance data of IR-range sensors. We could easily detect the region of a human body in an image by using the IR-range sensor because the sensor could obtain many 3D positions of measurement objects' surface. In addition, the IR-range sensor was robust to occlusion than an optical motion capture system in the prototype system. We, therefore, expected that we could use this method for investigations about cultural differences. In this section, we briefly describe the method.

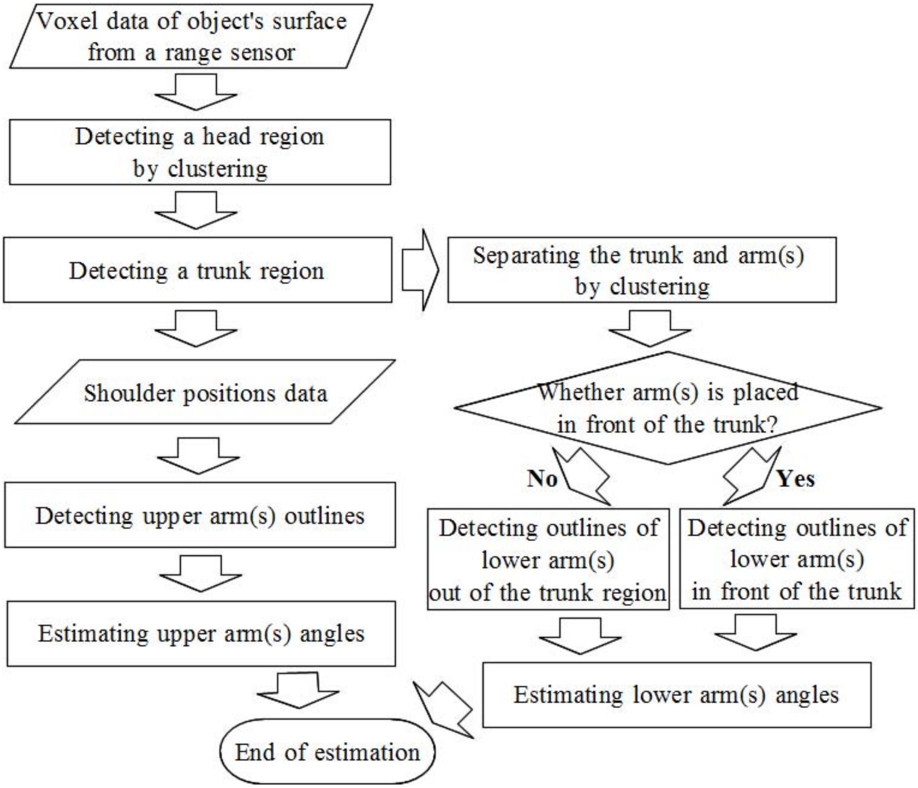

Fig. 4. The overview of the estimation of arm angles

The overview of the estimation process is shown in Figure 4. The input of the estimation process is voxel data of object's surface. In the first step, the horizontal widths of scattered parts of the measured object are used as the input features for a clustering technique to segment into head region and body region. In the next step, a trunk region is detected on the basis of structure of the body. From the third step, the estimation process diverges into two; estimation of upper arm angles and lower arm angles.

In the upper arms estimation, shoulder positions are firstly estimated from the trunk region on the basis of structure of the body. Secondly, the outline of the upper arms is detected by searching voxel data from the shoulder positions to the outside of the trunk. Finally, detected voxels of the upper arm outlines are used as the input features for a straight-line approximation of upper arm angles.

In the lower arms estimation, the distance of the voxel data is firstly used as the input features for a supervised clustering technique to segment into the trunk and the arms in the trunk region. When either arm is placed in front of the trunk, the outline of lower arms is detected by searching voxel data of the lower arm clusters. When both of the arms are placed out of the trunk region, the vertical widths of scattered parts of the measured object are used as the input features for a clustering technique to segment into the arms region out of the trunk. The outline of the lower arms is then detected by searching voxel data of the arm clusters out of the trunk region. Finally, detected voxels of the lower arm outlines are used as the input features for a straight-line approximation of lower arm angles.

(a) IR-range sensor data (b) arm region clustering (c) arm angles estimation

Fig. 5. An example of arm angles estimation

Figure 5 shows an example of an arm angles estimation; (a) shows the IR-range sensor data, (b) shows the result of arm region clustering when either arm is placed in front of the trunk and (c) shows the result of the arm angles estimation. In (b), the dark gray region indicates a right arm region. This means that we could correctly segment the trunk region and the arm region. In (c), two white lines indicate upper and lower arm angles. This means that we could correctly estimate the arm angles.

5 Discussions

5.1 The Concepts to Capture Gaze Directions, Facial Expressions and Hand Gestures

In this section, we discuss the concepts to capture gaze directions, facial expressions and hand gestures. We expect that these are needed for investigations about cultural differences.

Gaze direction. The first author developed a real-time system for measuring gaze directions without head restraint [11]. The system could measure gaze directions in a small space (30cm high, 30cm wide and 20cm depth) because stereo cameras could measure in the small space.

By using IR-range sensors, we can obtain many 3D positions and the rough texture of face in real time (Figure 6). We can thus improve the system with same accuracy (approximately ± 2 degrees) in larger space by using both IR-range sensors and cameras. The system can be used for CEBE.

Fig. 6. The examples of captured data of faces and hands by an IR-range sensor

Facial expressions. There are many software which convert facial features of a person into facial features of an avatar. We can easily invoke the external program as a plug-in from the virtual agent which we use because the virtual agent are worked on GECA Framework. When we use the software which can export facial features in real time, the virtual agent can express the facial expressions of the user in CEBE.

Hand gestures. We did not implement the method to capture hand gestures because we preferentially implemented the methods to capture human behavior to apply CEBE for using robots. It is difficult to capture hand gestures by using a motion capture system needed makers. We expect that we can capture hand gestures because we can obtain 3D positions of the surface by using IR-range sensors.

Ghobadi et al. [3] proposed a system to control hand based robot using IR-range sensors. Figure 6 also shows that IR-range sensors can capture the features of a hand. In addition, we can easily trace the hand positions because we estimate 3D-angles of arms. We thus expect that we can capture more detailed hand gestures by using both IR-range sensors and pan-tilt cameras.

5.2 Limitations

We will have to develop some implementations so that an agent reflects human behavior and can interact with people with cultural background.

It is most important for natural interaction to compensate for the delay time of agent actions. This delay was pronounced when a robot was controlled. To resolve this issue, we will have to implement a predictive control mechanism. On the other hand, we need to investigate the influence of the delay time of agent actions in encultured agent-human interaction.

It is also important to develop a better method to give feedback to a system user. In prototype system, we give feedback by using visual information. This method, however, is not effective when an ECA reflects the eye directions of the user. The user cannot perform appropriate actions without feedback. On the other hand, large feedback prevents the user from natural interaction. In addition, human behavior in cross-cultural communication is often expressed unconsciously. We have to consider whether we had better to give feedback or not.

5.3 Target Interactions

The prototype CEBE which we are implementing will be able to detect head direction, posture of the upper body and 3D angles of arms, and control a head and arms of a robot or a virtual agent based on the detected data. By using CEBE, we will be able to gather and analyze the information needed for realizing enculturated agent-human interaction. We can use the captured data for unsupervised simultaneous learning of gestures, actions and their associations, such as Embodied Interactive Control Architecture (EICA) [16].

6 Conclusion

6.1 Summary

In this paper, we proposed an environment in which a person can interact with ECAs controlled by captured behavior of another person with cultural background, and we can analyze the data obtained in actual enculturated agent-human interaction. In particular, we proposed a design and an implementation of the system, CEBE.

In future works, we will focus on a negotiation situation in which people have to express their opinions and speculate on the inner states of a communication partner, and we will conduct a preliminary experiment to confirm the effectiveness of CEBE in an enculturated agent-human interaction. In the experiment, we expect to reveal influences of cultural background in some situations, such as detecting the degree of agreement or disagreement, commitment to the communication, emotional representation, and how strong their assertion in the communication.

References

1. Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.): Embodied Conversational Agents. The MIT Press, Cambridge (2000)
2. Cassell, J., Tartaro, A., Oza, V., Rankin, Y., Tse, C.: Virtual Peers for Literacy Learning. Educational Technology. Special Issue on Pedagogical Agents XLVII, pp. 39–43 (2007)

3. Ghobadi, S., Loepprich, O., Ah-Madov, F., Bernshausen, J., Hartmann, K., Lof-Feld, O.: Real time hand based robot control using 2D/3D images. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Remagnino, P., Porikli, F., Peters, J., Klosowski, J., Arns, L., Chun, Y.K., Rhyne, T.-M., Monroe, L. (eds.) ISVC 2008, Part II. LNCS, vol. 5359, pp. 307–316. Springer, Heidelberg (2008)
4. Hirsch, M., Lanman, D., Holtzman, H., Raskar, R.: BiDi screen: a thin, depth-sensing LCD for 3D interaction using light fields. In: ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies: Adaptation, p. 62. ACM, New York (2009)
5. Huang, H.H., Cerekovic, A., Nakano, Y., Pandzic, I.S., Nishida, T.: The design of a generic framework for integrating ECA components. In: Padgham, L., Parkes, D., Muller, J.P. (eds.) The 7th international conference of autonomous agents and multiagent systems (AAMAS 2008), pp. 128–135 (2008)
6. Huang, H.H., Cerekovic, A., Pandzic, I.S., Nakano, Y., Nishida, T.: Enculturating human-computer interaction. AI and Society 24, 225–235 (2009)
7. Iacobelli, F., Cassell, J.: Ethnic identity and engagement in embodied conversational agents. In: International Conference on Intelligent Virtual Agents (IVA 2007), pp. 57–63 (2007)
8. Isbister, K.: Building bridges through the unspoken: embodied agents to facilitate intercultural communication. In: Agent culture: human-agent interaction in a multicultural world, pp. 233–244. Lawrence Erlbaum Associates, London (2004)
9. Nishida, T. (ed.): Conversational Informatics. An Engineering Approach. John Wiley and Sons, Chichester (2007)
10. Ohmoto, Y., Ohashi, H., Takahashi, A., Nishida, T.: Capture and Express Behavior Environment (CEBE) as WOZ system to realize effective human-agent interaction. In: Proceedings of International Workshop on Interacting with ECAs as Virtual Characters in AAMAS (2010) (in press)
11. Ohmoto, Y., Ueda, K., Ohno, T.: Real-time system for measuring gaze direction and facial features: Towards automatic discrimination of lies using diverse nonverbal information. AI and Society: The Journal of Human-Centred Systems 23(2), 187–200 (2009)
12. Prendinger, H., Ishizuka, M. (eds.): Life-Like Characters. Tools, Affective Functions, and Applications, Cognitive Technologies. Springer, Heidelberg (2004)
13. Rehm, M., Nakano, Y., André, E., Nishida, T.: Enculturating human-computer interaction. AI and Society 24, 209–211 (2009)
14. Rehm, M., Nakano, Y., André, E., Nishida, T., Bee, N., Endrass, B., Wissner, M., Lipi, A.A., Huang, H.H.: From observation to simulation: generating culture-specific behavior for interactive systems. AI and Society 24, 267–280 (2009)
15. Tachi, S., Kawakami, N., Nii, H., Watanabe, K., Minamizawa, K.: TELEsarPHONE: Mutual Telexistence Master-Slave Communication System based on Retroreflective Projection Technology. SICE Journal of Control, Measurement, and System integration 1(5), 1–10 (2008)
16. Mohammad, Y., Nishida, T., Okada, S.: Unsupervised Simultaneous Learning of Gestures, Actions and their Associations for Human-Robot Interaction. In: The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2537–2544 (2009)