3D Bow and Posture Measurements for Virtual Reality Customer Service Training System

Tomoya Furuno, Satoru Fujita, Wang Donghao, and Junichi Hoshino
University of Tsukuba, Ibaraki, Japan
Email: furuno.tomoya@entcomp.esys.tsukuba.ac.jp, jhoshino@esys.tsukuba.ac.jp

***Abstract*—Many service industries focus on reception (hospitality) and improving customer satisfaction by understanding customer needs and making good impressions. This kind of customer service is learned over a long period by manuals, role-playing, and on-the-job training. However, due to various limitations, e.g., time, it is difficult to repeatedly practice responding to different situations. Thus, in this paper, we propose a scenario-based virtual reality customer service training system that employs multimodal recognition of trainees. The proposed system measures bow and stoop degrees from skeletal posture information extracted using an RGB-D camera. Then, the proposed system provides feedback, i.e., visual aids and comments, to the trainee. We also investigate how multiple training sessions can change a trainee's customer service skills.**

***Index Terms*—customer service, scenario-base, training VR System, 3D posture**

I. Introduction

Traditionally, reception skills in customer service are learned over a long period through on-the-job training with guidance from senior employees and other individuals. However, in such cases, in addition to the psychological burden of failures, it is difficult to repeatedly exercise reception skills in different situations.

Muramoto *et al.* [1] developed a simulation game to train newly hired female clerks at women's fashion stores. This simulation software included customer dialog that reflects the interrelationships among customer, products, and the clerk. In addition, Robert *et al.* [2] developed and evaluated a character-based dialog training system to train informed consent skills. They demonstrated that test subjects who trained with a virtual character achieved better performance in actual interhuman interviews than those who trained only documents.

In this paper, we propose customer service training system that involves body movement measurement and user feedback. The proposed system allows a trainee playing a staff role to repeatedly practice listening to problems and evaluating potential customer needs via multimodal dialog with both speech and gestures.

The proposed system measures the bow and stoop degrees from skeletal information extracted using an RGB-D camera (i.e., a Microsoft Kinect device). The system then provides feedback to the trainee comprising visual aids and comments. We also investigated how multiple training sessions with the proposed system can change a trainee's customer service skills.

II. Related Research

In terms of Virtual Training (VR) systems for situations where physical response is important, Jaikyung *et al.* [3] produced a team-based firefighter training platform using VR, AR, and haptic technologies because it is difficult to train firefighters in practical fire situations. In addition, Yazhou *et al.* [4] produced an American football training system that employed a VR environment, and they evaluated its effectiveness over three days of training. The results demonstrated that the average training score improved by 30%. These studies are primarily immediate decisions and coping behaviors; however, in customer service, it is important to train employees to understand problems through conversation so they make take appropriate action.

In terms of customer service training, Otsuki *et al.* [5] developed a VR work training system that focuses on awareness and priority judgment in the restaurant service industry. With this system, it is possible to train employees to be aware of what is occurring at multiple locations and determine appropriate priorities by placing customers at multiple tables at different times. In this study, the evaluation of customer service behavior is assessed quantitatively by recording the time required for a given task. Oyanagi *et al.* [6] constructed a training simulator to develop psychological skills for the service industry using a VR environment. Here, users in remote locations played the roles of teacher and student, and they interacted with each other in VR space over a network. In these studies, the instructor evaluated the trainee's customer service behaviors.

III. Overview of the Customer Service Training System

Fig. 1 shows the multimodal dialog process involved in customer service. In customer service, as discussed in various reception studies, it is important to be able to read the other person through observation, listening, and nonverbal communication, e.g., facial expressions, line of sight, and gestures. The necessary reception skills are summarized in the following.

Manuscript received August 10, 2021; revised November 17, 2021.

doi: 10.18178/joig.9.4.152-156

Mental attitude: A way of thinking in which customers are treated with sincerity.

Understanding the customer: Recognition of the customer's situation according to their age, gender, appearance, behavior, and speech.

Response based on expertise: This involves predicting appropriate responses and providing convincing suggestions based on specialized business knowledge.

Figure 1. Multimodal dialogue process of customer service.

The service training system must have the following characteristics.

1) Easy to adapt to different training scenarios

To train effective reception skills, various scenarios should be available, including scenarios that require listening to the customer's situation and needs, business knowledge, and emotional sensitivity.

2) Recognition of multimodal dialog with the customer agent

Customer agents must be able to recognize the trainee's voice and gestures, respond to them, and exhibit emotional reactions.

3) Evaluation of the interaction

The system must be able to evaluate whether convincing responses were given based on appropriate business knowledge.

4) Feedback to the trainee

Effective feedback should be provided to the trainee to guide improvement.

In the current implementation, a VR environment model with a customer character was implemented using Unity (version 2019.1). Here speech recognition was achieved via Windows.speech, and the Python Transitions package was used for finite state machine-based dialog control. In addition, eye tracking was realized using the HTC VIVE Pro, and gestures were recognized via bone extraction using a Kinect RGB-D camera to classify different types of bows (i.e., slight bow, salute, and respectful bow).

IV. Bow and Posture Measurements

We measured the bow and stoop degrees by extracting skeletal information using the Kinect RGB-D camera.

Bowing at the correct angle is important to provide a good impression to customers. Here, bows were divided into nod, salute, most salute in Japan. Table I and Fig. 2 summarize each type of bow [7], [8].

TABLE I. Types of Bow

Type	Degree	Proper situation
Nod	15	Usual greeting
Salute	30	Showing respect
Most salute	45	Apologies and deep thanks

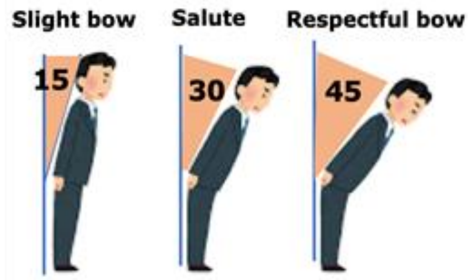

Figure 2. Different types of bow (slight bow, salute, respectful bow).

Nods are used in the calling phase, the highest salute is used in the problem confirmation and problem-solving phase, and the salute is used in the see-off phase. We calculated the angle of the user's bow by calculating the angles of two vectors obtained from the skeletal information (Fig. 3) using the Kinect, i.e., the vectors from SpineShoulder at normal times and when bowing to neck.

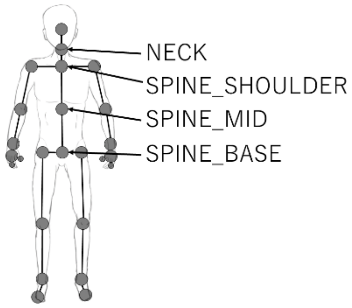

Figure 3. Skelton Information using Kinect.

A stoop is a posture where the neck is placed to the front slightly and the back is rounded. In this posture, the curvature of the spine and the shoulder often move forward [9]. In customer service, it is not preferable to respond in a stooped state; therefore, effective customer service skills should include "keeping the correct posture" during all phases.

SpineShoulder, SpineBase, and SpineMid were used in Fig. 3. Here, the user's stoop is calculated and evaluated.

$$\boldsymbol{D}_i = \left\{(\boldsymbol{b}_i - \boldsymbol{a}_i) \cdot \frac{(\boldsymbol{b}_i - \boldsymbol{c}_i)}{|\boldsymbol{b}_i - \boldsymbol{a}_i|^2}\right\}(\boldsymbol{b}_i - \boldsymbol{a}_i) - \boldsymbol{c}_i \quad (1)$$

by calculating the distance *Di* between a straight line and SpineMid point, which connects the SpineShoulder and SpineBase points. If the coordinates of SpineShoulder, SpineBase, and SpineMid are denoted ai, bi, and ci, respectively, *Di* can be expressed by Eq. (1).

To determine a threshold, we measured the variation in *Di* for six subjects (Table II). Here, the distance *Di* value was measured for 5 s for both up-and-down conditions and stoops, and the average was calculated. Considering fluctuation in the *Di* value due to slight changes in posture, the threshold of *Di* was set to 0.1 m.

TABLE II. MEASUREMENT OF DIFFERENCE BETWEEN STANDING STRAIGHT AND STOOP

Subject	Standing(m)	Stoop (m)	Difference
1	0.27	0.12	0.15
2	0.25	0.02	0.23
3	0.49	0.34	0.15
4	0.35	0.24	0.11
5	0.31	0.20	0.11
6	0.39	0.32	0.07
Average	0.34	0.21	0.14

V. USER FEEDBACK METHOD

Feedback is required for users to improve their performance over repeated customer service training sessions.

- If the trainee's bow is appropriate, the error in the angle of the bow versus the angle of an appropriate bow is displayed.
- If the trainee's bow is inappropriate or if the trainee failed to bow at the appropriate moment, the appropriate type of bow, and its angle are displayed.
- If the trainee's posture becomes a stoop, feedback is displayed to identify this out (Fig. 4).

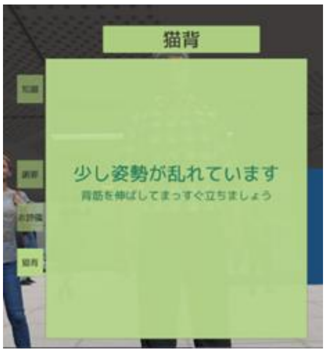

Figure 4. Feedback of stoop (posture should be a little more straight).

VI. SCENARIO-BACED TRAINING

In the proposed system, a customer service role play scenario with a customer character is conducted according to a specific scenario, i.e., complaint correspondence training. Skills to train are commonly referenced in customer service literature [10]-[16], we select "check for problems," "apology," "problem-solving," "wording," "handling knowledge," and "posture."

In this scenario, the user responds to a customer complaint as a member of the ground staff at an airport. Note that the complaint was created in reference to the actual rules and cases of an airline to approximate a real-world business situation. These scenarios consider several cases, e.g., damaged luggage and refunding airline tickets.

The scenario involved three utterances for the customer character and four utterances for the ground staff (the user). The actions corresponding to each utterance phase are shown in Table III. To avoid a breakdown in the conversation, the conversation was led by the customer, and utterances were only made in the user's utterance phase.

In each speech phase of the ground staff, the user thinks of specific speech content and speaks it. In response to the four speech phases of the user, the customer has three speech phases, i.e., the "problem presentation," "response request," and "exit" phases. In each phase (except for the "exit" phase), the information required for the user's speech is presented by customer's speech. For example, the customer describes the current situation in the "problem presentation" phase.

To allow beginners to learn and train, the procedures to be explained, customer service knowledge, and the direction of the utterances are set for each scenario, and the user can check them as hints in the system.

TABLE III. SCENARIO CONFIGURATION

	Speaker	Functions	Utterances
1	User	Call	talk to a customer in trouble
2	Customer	Problem Presentation	ask a question about a problem
3	User	Check for problems Apology	answer the customer's question
4	Customer	Response request	make a claim to a problem
5	User	Problem solving / Apology	solve problems and give apology
6	Customer	Exit	satisfied with the response, leaving
7	User	See off	add a word of see-off to a customer

VII. USER STUDY

A. *Experimental Setup*

An evaluation experiment was conducted to verify the effectiveness of claim response training of the proposed system. In this experiment, after receiving an explanation of simple customer service knowledge, we trained the users to respond to complaints using in-person role play scenario and the proposed system for scenarios with the same configurations. The participants included nine university and graduate students (seven males and two females with an average age of 22.5 and standard deviation of 1.21). Note that six participants were experienced customer service students. After training all participants, two independent evaluators watched a video of the training and provided an evaluation via a questionnaire of five methods. The questionnaire is shown in Table IV.

The experimental procedure is described in detail in the following.

TABLE IV. QUESTIONNAIRE FOR VIDEO EVALUATION (FIVE POINTS SCALE)

1	Do you think the subject's standing posture is maintained correctly?
2	Do you think the angle of the subject's greeting bow (15, degrees) is correct?
3	Do you think the angle of the subject's apology bow (45, degrees) is correct?
4	Do you think the angle of the subject's bow (30 degrees) is correct?

1) *Manual learning*

A seven-minute learning phase was conducted using a written manual (four pages) prepared by the authors. This learning phase was performed to provide the participants with the minimum required customer service knowledge and the knowledge required to perform the interpersonal role play scenario.

2) Interpersonal role play scenario (before the system training)

The customer service role play scenario prepared by the authors was performed with a performer, and the scene was filmed. Note that the structure of the role play scenario was the same as that of the training system; however, the specific content differed from that used in the training system.

3) System training

Participants wore head mounted display and performed the role play session with a customer character in the virtual space of a customer service scene. The role play was conducted according to four different scenarios created by the authors. As shown in Fig. 5, one training task comprises three phases, i.e., the practice, test, and rest phases, and the user experienced the training task three times. In the practice and test phases, the training was conducted in the scenario. In the training phase, hints for appropriate responses were displayed. However, these hints were not displayed in the test phase. In addition, the test phase was filmed.

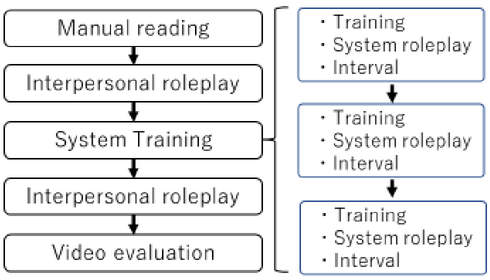

Figure 5. Flow of experiment.

4) Interpersonal role play (after the system training)

Here, a seven-minute learning session was performed using the same manual used in manual learning. Then, an interpersonal role play scenario was performed with the performer as in 2), and the scene was filmed. Note that the role play scenario was the same as that experienced in Interpersonal role play (before the system).

5) Video evaluation

After the interpersonal role play scenario was performed, four evaluators who were not involved in the experiment, watched the recorded video and evaluated the proposed system using a questionnaire (Table III). To emphasize the impressions given for ordinary customers, the evaluators were not customer service experts.

VIII. Results and Discussion

In this paper, the significance level is indicated as $p<0.05$ for * and $p<0.01$ for **.

A. System Role Play

We obtained data from the nine subjects and applied the Friedman's test to the first, second, and third training sessions of the role play system.

1) Training bow scores

Table V shows that we observed significant differences in the items related to bowing (items 2, 3, and 4), and the mean values are on the increase. Thus, we believe that that proposed system helped the participants learn to select an appropriate bow according to the given situation.

TABLE V. System Role Play Evaluation Score

	System role play evaluation Average Score			
	First (SD)	Second (SD)	Third (SD)	p-value
1	3.47(1.0)	3.71(0.97)	3.62(0.96)	0.28
2	3.60(1.0)	3.93(1.1)	3.98(1.1)	0.0086**
3	3.49(0.94)	4.11(0.83)	4.07(0.84)	8.2e-05**
4	3.89(0.83)	4.16(0.71)	4.24(0.68)	0.0033**

2) Training posture scores

For the posture scores, no significant differences were observed in the relevant item. We believe that this finding was caused by the fact that the participants were aware their posture was being measured by the proposed system; thus, the participants were conscious of maintaining good posture.

B. Interpersonal Role Play

For the in-person role play scenario, the Wilcoxon signed-rank test was performed before and after the participants used the proposed system. The results are shown in Table VI. As can be seen, the score for items related to bows (i.e., items 2, 3, and 4) for the second training session were significantly higher than those for the first training session. Thus, we believe that the skills learned by using the system were demonstrated by the participants.

For posture (item 2), no significant difference was observed in the system training; however, a significant difference was observed in the interpersonal role play. Here, we believe that the participants were not very conscious of their posture in the first interpersonal role play but became more conscious of posture through the feedback during training. Thus, the participants were able to maintain good posture in the interpersonal role play after using the proposed system.

The items related to bowing and posture demonstrate that the skills acquired using the proposed systems are effective because the scores were significantly improved in-person.

These results demonstrate that it is possible to learn appropriate bowing and posture behaviors using the proposed training system. In addition, the skills acquired by using the proposed system can be used effectively when dealing with other people.

TABLE VI. Interpersonal Role Play Evaluation Score

	Interpersonal role play evaluation Average Score		
	First (SD)	Second (SD)	p-value
1	2.84(1.1)	3.53(1.2)	0.00050**
2	2.84(1.2)	3.24(1.2)	0.020*
3	2.38(1.0)	3.40(1.1)	3.7e-05**
4	2.49(1.2)	3.44(1.0)	0.0013**

IX. Conclusion

In this paper, we have proposed a VR system to perform claim response training using interactive characters in VR space. The results of an evaluation experiment demonstrated that it is possible to acquire appropriate bowing and posture behaviors using the proposed training system. In addition, it is clear that the skills acquired would be effective in-person.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Tomoya Furuno primarily conducted the research; Satoru Fujita conducted user study; Wang Donghao created the scenario; Junichi Hoshino supervised and produced the research.

Acknowledgment

This work was partly supported by Council for Science, Technology and Innovation, "Cross-ministerial Strategic Innovation Promotion Program (SIP), Big-data and AI-enabled Cyberspace Technologies". (Funding Agency: NEDO).

References

[1] R. Muramoto, T. Kaneda and J. Tanabe, "A simulation game software for training use OF sales-talking in fishion shop," The Special Interest Group Technical Reports of IPSJ, vol. 95, no. 23, pp. 59-65, 1995.

[2] R. C. Hubal and R. S. Day, "Informed consent procedures: An experimental test using a virtual character in a dialog systems training application," *Journal of Biomedical Informatics*, vol. 39, no. 5, pp. 532-540, 2006.

[3] J. Lee, M. Cha, B. Choi, and T. Kim, "A team-based firefighter training platform using the virtual environment," in *Proc. VRCAI*, December 2010.

[4] Y. Huang, L. Churches, and B. Reilly, "A case study on virtual reality American football training," in *Proc. Virtual Reality International Conference*, April 2015.

[5] M. Otsuki and T. Okuma, "Development of a VR business training system for the restaurant service industry," IPSJ SIG Technical Report, vol. 2020-HCI-187, no. 3, pp. 1-6, 2020.

[6] A. Oyanagi, *et al.*, "Training simulator for service industry using virtual reality environment," *Transactions of the Virtual Reality Society of Japan*, vol. 25, no. 1, pp. 78-85, 2020.

[7] H. Koga, "A study of bowing: Four types of bow," Kaetsu University Research Papers, vol. 55, pp. 57-71, 2012.

[8] S. Eikou, "Analysis of the relation between the bow and customer satisfaction in concierge service," Lecture Papers at the Autumn Meeting of the Society of Precision Engineering, pp. 583-584, 2010.

[9] Improve posture by correcting the winding shoulder. [Online]. Available: http://www.hiranogh.com/hirano/medical_department/News/rihabili-news007.html

[10] Y. Nagao, "Illustration of impressive customer services," *Shuwa System*, 2018.

[11] Complaint Handling Manual. [Online]. Available: http://www.business-sol.jp/category/1505554.html?page=2

[12] Communication to convey emotional understanding. [Online]. Available: https://www.insource.co.jp/mailmagazine/clam20120626.html

[13] Language and way of talking when responding to complaints. [Online]. Available: https://www.insource.co.jp/mailmagazine/clam20120605.html

[14] It's different at the angle! Types of Bows 3 and The Meaning of Hospitality Manners. [Online]. Available: https://careerpark.jp/692

[15] The type of bow and the angle of each are summarized! With the point of hospitality manners. [Online]. Available: https://and-plus.net/ojigi/

[16] What is Stoop postures. [Online]. Available: https://roppongi-tetote.com/neko.php

Tomoya Furuno is a graduate student of University of Tsukuba. His research interests include entertainment computing and scenario-based VR system. He is working on 3D posture measurement and application for customer service training, and published a paper at IPSJ (Information Processing Society in Japan).

Satoru Fujita is an alumni of University of Tsukuba. His research interests include entertainment computing and game technologies. His master thesis was on the scenario-based customer service training system. He was a main architect and programmer of the training VR system. He is currently working at a game company as an engineer.

Wang Donghao is an alumni of University of Tsukuba. His research interests include entertainment computing and game technologies. His master thesis was on the scenario-based customer service training system. He was working on creating service training scenarios and designed user studies. He is currently working at a game company as a concept designer.

Junichi Hoshino is an associate professor of University of Tsukuba. His research interests include entertainment computing, game technologies, and digital storytelling. He is currently a vice chair of IFIP TC14 Entertainment Computing Technical Committee, and a chair of WG14.4 Entertainment Games. For more information please see www.entcomp.esys.tsukuba.ac.jp.