



**HAL**  
open science

# Customer Service Training VR Game System Using a Multimodal Conversational Agent

Tomoya Furuno, Yuji Omi, Satoru Fujita, Wang Donghao, Junichi Hoshino

► **To cite this version:**

Tomoya Furuno, Yuji Omi, Satoru Fujita, Wang Donghao, Junichi Hoshino. Customer Service Training VR Game System Using a Multimodal Conversational Agent. 19th International Conference on Entertainment Computing (ICEC), Nov 2020, Xi'an, China. pp.277-281, 10.1007/978-3-030-65736-9\_25 . hal-03686022

**HAL Id: hal-03686022**

**<https://inria.hal.science/hal-03686022>**

Submitted on 2 Jun 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Customer Service Training VR Game System Using a Multimodal Conversational Agent

Tomoya Furuno<sup>1</sup>, Yuji Omi<sup>1</sup>, Satoru Fujita<sup>1</sup>, Wang Donghao<sup>1</sup>, and Junichi Hoshino<sup>1</sup>

<sup>1</sup> University of Tsukuba, Ibaraki, Japan  
furuno.tomoya@entcomp.esys.tsukuba.ac.jp, jhoshino@esys.tsukuba.ac.jp

**Abstract.** Hospitality is the primary focus in many service industries. Staff can improve customer satisfaction by grasping the customer's situation accurately and responding to it appropriately. To cultivate these skills, we propose a VR training system based on multimodal recognition of the customer's status during service. A preliminary user study of a claim handling situation is also summarized.

**Keywords:** VR training system, customer servers, hospitality

## 1 Introduction

Traditionally, customer service skills are learned over a long time through experience on the job and supervision from senior employees and other relevant people. However, this approach involves experiencing failures due to lack of experience, thereby creating a psychological burden on the staff. Even when on-the-job training is provided, exercising many situations repeatedly is difficult. Muramoto et al. [1] developed a customer service training software for newly hired female clerks at women's fashion stores. The software simulated customer dialog and the interrelationships among the customer, product, and clerk. Robert et al. [2] developed a character-based dialog training system for informed consent skills and evaluated its effectiveness. The study demonstrated that the subjects who trained with virtual characters achieved better performance in an interview with a real person than the subjects who trained using only written instructions.

In the present work, we describe both the concept and current prototype of a customer service VR game system using a multimodal conversational agent with dialog involving speech and gestures. This system enables the trainee to play the role of staff in various situations as many times as needed to learn to not only understand the customers' problems and anticipate their potential needs but also provide accurate solutions and appropriate psychological care.

## 2 Overview of Customer Service Training System

High-quality customer service creates a positive impression of the stores and products leading to an increase in the total number of visitors and repeated purchases. If a clerk leaves customers unsatisfied, they are unlikely to return even if the quality of the products is good. Skillful hospitality is especially important when customer experience is an integral part of the service, for example, in airports, hotels, and restaurants.

The following competencies are considered to be important.

- 1) **Understanding potential needs:** estimating problems from the customers' characteristics and behavior as well as the current situation. For example, if a customer appears confused while standing in front of an airport ticketing machine, he or she may not know how to operate it.
- 2) **Proposing solutions:** being able to present an appropriate solution based on available information. For example, if there is a problem with an airplane, the solution must take time constraints and other circumstances into account.
- 3) **Creating a positive experience:** social cues of staff such as facial expressions, tone of voice, and gestures are also important for leaving a good impression.

A skilled customer service staff member should be able to

- deal with all situations described in the manual;
- accommodate a large number of customers in a limited amount of time;
- handle claims;
- and improvise appropriate responses to rare and unique situations.

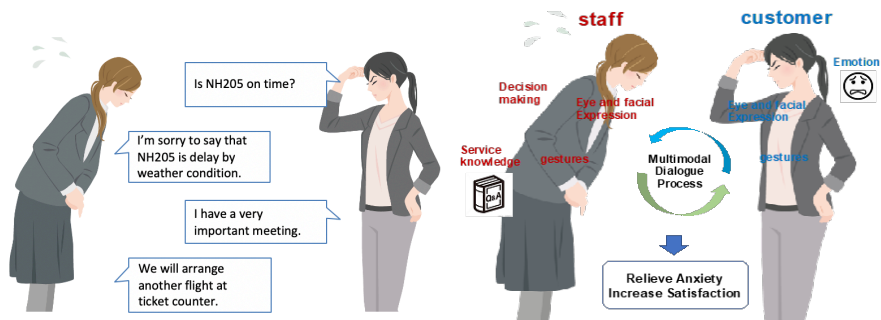
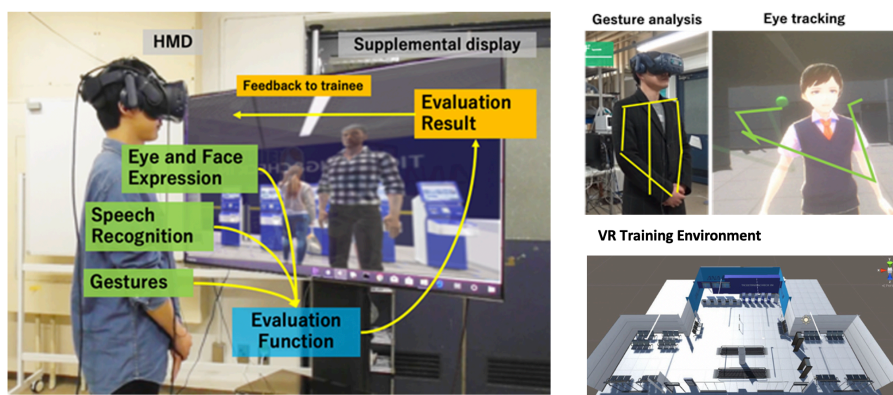


Fig. 1. Multimodal dialogue process of customer service

For the training to be successful, it is necessary to be able to recreate the details of relevant scenarios, such as time, place, occasion, and customer characteristics. Training based on specific scenarios includes the goal-based scenario model in experiential education, content-based instruction in second language education, task-centered teaching, and simulation education in medical fields.

We built our system to conform to the following requirements:

- It should be able to simulate various training scenarios, including those that require listening to the customers and understanding their needs, require business knowledge, and require emotional care.
- The customer agents should be able to proceed appropriately by analyzing the voice and gestures of the trainee and responding to them verbally and through showing emotional reactions.
- It should be able to evaluate whether the trainee has provided convincing responses based on business knowledge.
- It should be able to provide feedback and suggestions for improvement to the trainee.



**Fig. 2.** Concept of customer service VR training system using conversational agent

The VR environment and customer character were implemented using Unity version 2019.1. Speech recognition was done using Windows Speech Recognition, and Transitions, a Python package, was used for the finite-state-machine-based dialog. Eye tracking was done using HTC VIVE Pro, and gestures were recognized using bone extraction facilitated by an RGB-D camera (Microsoft Kinect). Different types of bows (slight bow, salute, respectful bow) were distinguished.

### 3 Airport Ground Staff Scenarios

While hospitality is crucial for a variety of services, in the present work we focus on the ground staff of airline companies serving customers at airports. The role of the ground staff is to provide efficient assistance according to the airport and airline operation rules, ensuring the customer has a comfortable trip to their travel destination. The main work areas are around departure gates, check-in counters, and baggage claims. In recent years, areas around automatic boarding registration terminals and automatic baggage gates have become common as well. A sample scenario can be found below.

Customer: Excuse me, I would like to ask you a question.

Staff: Yes, how may I help you?

Customer: Is the flight A223 departing as planned?

Staff: I am sorry, it will be delayed by about two hours due to the typhoon.

Customer: Two hours late? I have plans immediately after arrival. Would it be possible to switch to another flight?

Staff: I am very sorry, other flights are all full. We apologize and provide a meal coupon to all passengers on flight A223, which you can use at this airport.

Customer: Could you arrange for a taxi upon my arrival?

Staff: Yes, would you like me to go ahead and do that?

Customer: Yes, please. Would the airline cover the fare in this case?

Staff: I am sorry, if the delay is due to a typhoon, the taxi fare will not be covered.

Customer: Okay, I understand.

## 4 Preliminary User Study

Four airport service beginners aged 20 to 27 years were asked to go through the training scenarios in the VR environment multiple times, and the changes in each person's responses with each repetition were analyzed.

The automatic check-in machines and baggage drops in airports were reproduced in the VR environment. The customer agent recognized the utterances of the staff in training using speech recognition and responded with speech synthesis and changes in facial expression.

The trainees played the role of ground staff, listened to the customer character complaining about the delay of the flight, and responded by offering a refund, providing assistance with transportation, etc. The trainees were asked to go through three sets of scenarios with 30-minute breaks in between. Each set comprised four scenarios with different variations, such as the reason for the delay (e.g., weather or aircraft failure), length of the delay, and arrival time. To analyze the trainee behavior, content of the conversations, their lengths, the time elapsed before the conversation starts, and lines of sight of the trainees were measured, and videos of their performances were taken.

We observed a tendency toward shorter delays before starting the conversations, accelerated decision-making, and a smoother manner of speech in all four subjects as the scenarios were repeated. Moreover, it was confirmed that the trainees were able to learn the necessary skills required to respond to the customers using this system repeatedly despite starting without preliminary knowledge. We also observed some voluntary actions such as the use of honorifics, other speech choices, and bowing.

## 5 Conclusion

We have described our current progress toward a customer service training system using a multimodal conversational agent. Future work will include long-term user studies with various training scenarios in business environments.

## **6 Acknowledgement**

This work was 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, J.Tanabe. A simulation game software for training use of sales-talking in fishion shop. The Special Interest Group Technical Reports of IPSJ. 1994-ICS-099, 59-65, 1994
2. C.Hubal, S.Day: Informed consent procedures: An experimental test using a virtual character in a dialog systems training application. Journal of Biomedical Informatics Volume 39, Issue 5, Pages 532-540, 2006