



Simultaneous Dialog Robot System

Takuya Iwamoto^(✉), Kohtaro Nishi, and Taishi Unokuchi

Cyberagent, Inc., Abema Towers 40-1 Udagawacho, Shibuya-ku,
Tokyo 153-0042, Japan
{iwamoto_takuya_xa, kohtaro_nishi,
taishi_unokuchi}@cyberagent.co.jp

Abstract. Various types of communication robots have been used in practical applications. Cases where robots are utilized as receptionist, among others, are increasing. However, this trend is not new nowadays. Generally, robots that are used for reception and customer services are developed on the assumption of one-on-one customer service. Thus, when one user interacts with the robot, other customers have to wait for their turn. In this research, we propose a system that can simultaneously serve many people by using one robot with multiple directional microphones and speakers.

Keywords: Robot · Humanoid · Information · Hospitality

1 Introduction

Currently, Amazon and Jindong operate as unmanned stores [1, 2]. In these stores, a user can do smooth shopping without using an accounting system. Henn-na Hotel [3] and Hama Sushi [4] in Japan have introduced a robot that allows users to experience hospitality. These robots are connected via a network to access big data in real time, answer user questions, and perform tasks.

Currently, customer service robots installed in department stores and information centers become common [5]. However, these robots can only speak to one user at a time, and people who want to talk/use the robot have to wait for their turn. Therefore, in this research, we develop a robot system wherein directional speakers are installed in the robot and can be set in pairs. Then, we describe its use case.

2 General System

This system can provide customer service not only from the front but also from the direction where the microphone and directional speaker are installed (Fig. 1). Through this directional speaker, the robot speaks at the same time. However, the user can hear only the voice directed to him/her, thereby enabling a one-to-many communication.

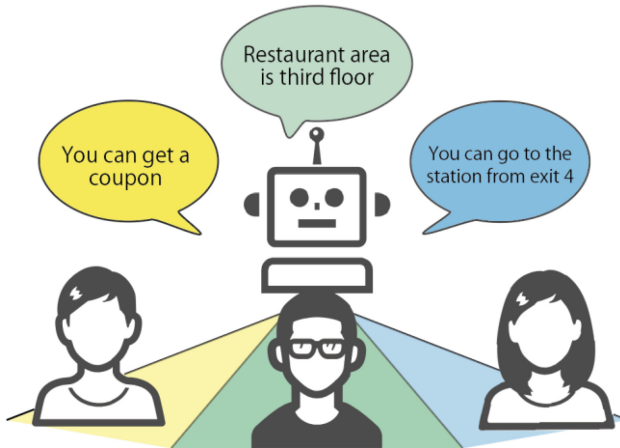


Fig. 1. System overall

3 System Components

Figure 2 shows the configuration of this system. The components of this system are a robot, two microphones, two directional speakers, and a personal computer (PC). The robot uses Pepper, which is a model for corporate use. Two USB microphones are attached on the front shoulders of Pepper. The directional speakers (Tristate, parametric speaker experiment kit) are fixed on to the two tripods behind the back of Pepper and are then attached to its head. The height was adjusted to ensure that the sound was heard from a close position (Fig. 3).

First, the microphone and speaker were connected to the PC. Second, the voice recorded from the microphone was sent to DialogFlow for voice recognition. Third, the answer registered in advance was collected, and the acquired character string was subjected to speech synthesis using Google Cloud text-to-speech application programming interface (API), which is a speech synthesis software as service. Finally, the answers to the questions were speech synthesized using Google’s Cloud text-to-speech API, and the output was released from the speaker.

The microphone and speaker were installed separately on the left and right, respectively, and the answer to the voice spoken on the left microphone is heard from the left speaker (the same for the right). This setup works in parallel.

4 Exhibition and Use Cases

This system was demonstrated and presented at Yahoo! JAPAN Hack Day 2018 which was held at Akihabara UDX on December 15–16, 2018. A user stated that “it can be used in various situations” and “waiting for practical use.” Hence, we considered the use case scenario of this system based on our experience at and opinion on the exhibition.

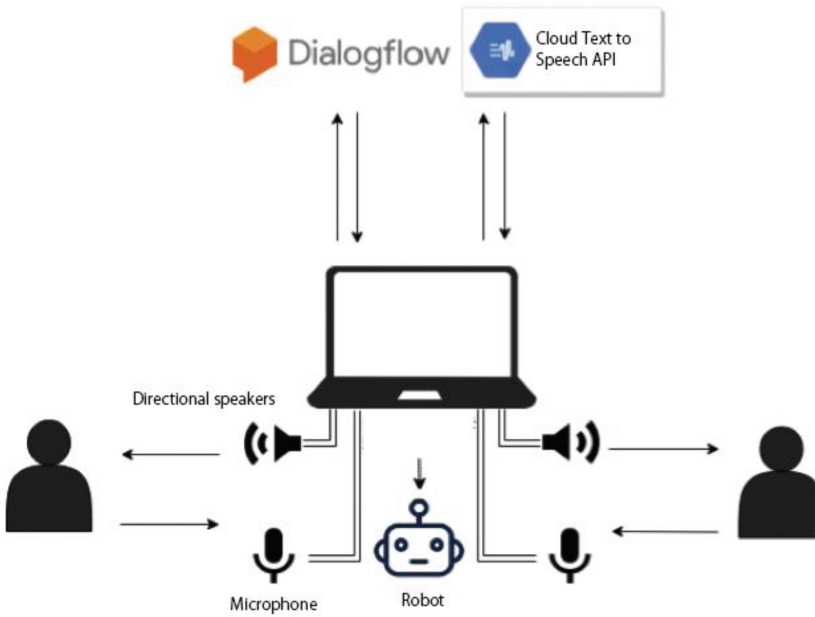


Fig. 2. System components

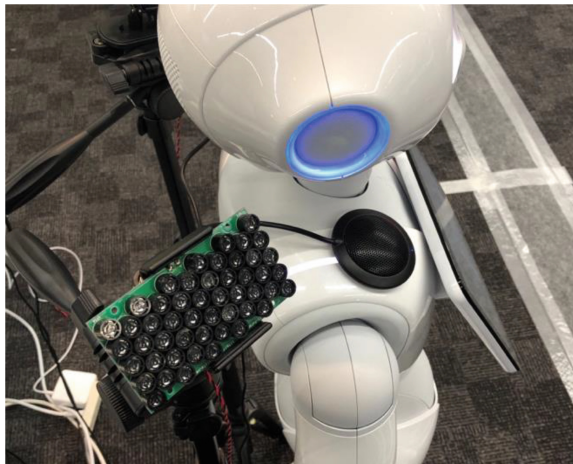


Fig. 3. Installation of the microphone and directional speakers

4.1 Use Case 1: Robot System that Can Simultaneously Speak Multiple Languages

Robot and guidance systems often support multiple languages in places with many visiting foreigners. Many of these systems allow users to change the language settings. The voice recognition function is also in progress, and Google Homepage has a function that can distinguish two languages set beforehand [6].

When utilizing this system, the user can use the system in the language that he/she prefers without setting the language by deciding the languages that can talk according to the position (Fig. 4).

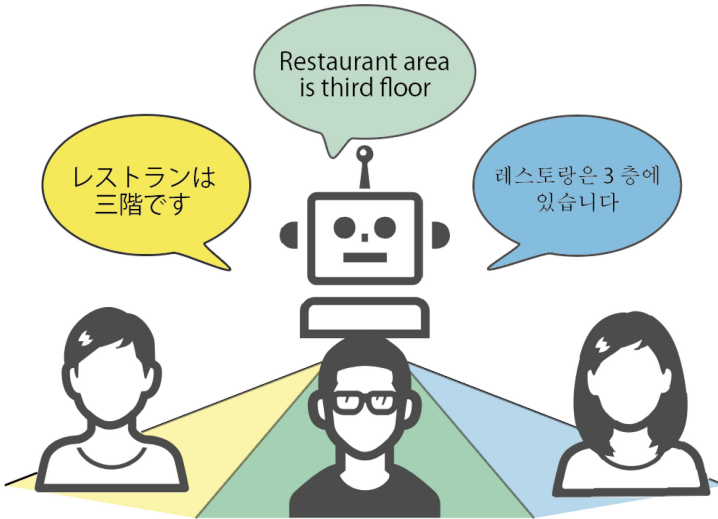


Fig. 4. Robot system that simultaneously speaks multiple languages

4.2 Use Case 2: Individual Approaching System

Robots approach pedestrians with voice and motion. By employing an omnidirectional speaker, the same message can be delivered to everyone in a room. When a robot announces an advertisement, such as introducing a product, only a few people need this message. Therefore, adding a function that can estimate personal attributes using the camera of this system and rendering a directional speaker movable can be useful, enabling an approach similar to the pedestrian attributes (Fig. 5).

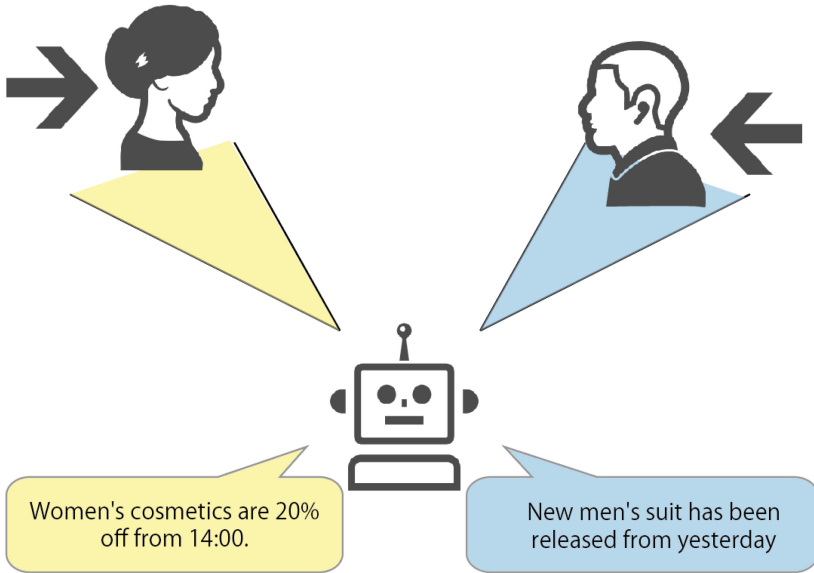


Fig. 5. Schematic of the Individual Approaching System

5 Future Work

In this study, we proposed a customer service robot by using multi-directional speakers. In the future, we will design an interaction that allows a user to feel that the system or robot is “speaking to her or him” only despite the robot speaks simultaneously to multiple people. Moreover, we will examine the feasibility of the use case scenario discussed in this paper.

References

1. Amazon Go. <https://www.amazon.com/b?ie=UTF8&node=16008589011>. Accessed 27 Mar 2019
2. Jian24. <http://www.jian24.com/>. Accessed 27 Mar 2019
3. Hen-na Hotel. <https://www.hennahotel.com/en/>. Accessed 27 Mar 2019
4. Hama Sushi Roboto. <https://www.softbank.jp/robot/biz/case/hamazushi/>. Accessed 27 Mar 2019
5. Watanabe, M., Ogawa, K., Ishiguro, H.: Field study: can androids be a social entity in the real world? In: Human-Robot Interaction (HRI), pp. 316–317 (2014)
6. Talk to the Google Assistant in multiple languages. <https://support.google.com/googlehome/answer/7550584?hl=en>. Accessed 27 Mar 2019