Developing an Alexa skill for learning sighted guide techniques

Iwao Sekita [Japan][1], Nobukiyo Sugisaki [Japan][1], Takuma Murakami [Japan][2], Mitsuhide Ishikawa [Japan][3], Tomoko Sakai [Japan][2], Yasuhiko Chiba [Japan][4], Noriko Yamaguchi [Japan][5], Ritsuko Endo [Japan][2], Hisako Shishido [Japan][2], Hisayuki Tatsumi [Japan][1]

Tsukuba University of Technology[1], NPO Shirogame[2], Tokyo Independent Living Support Center for the Visually Impaired[3], Miyagi Prefectural School for the Visually Impaired[4], Kansai Home for the Blind Woman[5]

Background

A Java application for learning sighted guide techniques was developed in 2006 (IMC12), and the relevant Moodle contents were developed in 2015 (basic version, IMC15) and 2017 (guide dog version, IMC16). The techniques are based on Shirogame's method and are applicable for a person with long-term as well as recent visual impairment. In the previous version, users were required to know computer security or have a login account and access to a mouse for operation. Recently, virtual assistants that are capable of voice interaction through smart speakers have become available, such as Alexa.

Methods

The built-in capabilities of Alexa are called skills. We have developed an Alexa skill for learning basic sighted guide techniques with 11 settings: basic guiding posture, passing through a narrow aisle, taking a seat on a chair, passing through doors, ascending and descending stairs, using an escalator and elevator, using a public toilet, and accessing a taxi, bus, or train. Each setting has several scenes, and the user selects one setting and one scene through voice command for instructions on sighted guide techniques. Each scene has several questions. Alexa asks the first question through text accompanied by an image. After the user selects an answer via voice, the user receives audio and visual feedback with a text, image, or video. The user says "Repeat" to obtain feedback again or "Next" to receive the next question. The command "Cancel" returns the user to the upper layer stage like scene selection stage.

Results and Discussion

The user can learn the techniques via voice command without mouse operation, thereby providing ease of access. Comments by test users indicated that the images, videos, and voice interactions helped them understand the techniques. This application is expected to help visually impaired persons travel with more freedom and safety.