International Journal of Informatics Society

10/23 Vol.15 No.2 ISSN 1883-4566

Informatics Society

Aims and Scope

The purpose of this journal is to provide an open forum to publish high quality research papers in the areas of informatics and related fields to promote the exchange of research ideas, experiences and results.

Informatics is the systematic study of Information and the application of research methods to study Information systems and services. It deals primarily with human aspects of information, such as its qu ality and value as a resource. Informatics also referred to as Information science, studies t he structure, algorithms, behavior, and interactions of natural and a rtificial systems that store, process, access and communicate information. It also develops its own conceptual and theoretical foundations and utilizes foundations developed in other fields. The advent of computers, its ubiquity and ease to use has led to th e study of info rmatics that has computational, cognitive and social aspects, including study of the social impact of information technologies.

The characteristic of informatics' context is amalgamation of technologies. For creating an informatics product, it is necessary to integrate many technologies, such as mathematics, linguistics, engineering and other emerging new fields.

Guest Editor's Message

Yoshiaki Terashima
Guest Editor of the Forty-fourth Issue of the International Journal of Informatics Society

We are delighted to have the Fortieth issue of the International Journal of Informatics Society (IJIS) published. This issue includes selected papers from the Sixteenth International Workshop on Informatics (IWIN2022), which was held online, August 31-September 3, 2022. The workshop was the sixteenth event for the Informatics Society, and was intended to bring together researchers and practitioners to share and exchange their experiences, discuss challenges and present original ideas in all aspects of informatics and computer networks. In the workshop, 25 papers were presented in eight technical sessions. The workshop was successfully finished with precious experiences provided to the participants. It highlighted the latest research results in the area of informatics and its applications that include networking, mobile ubiquitous systems, data analytics, business systems, education systems, design methodology, intelligent systems, groupware, and social systems.

Each paper submitted to IWIN2022 was reviewed in terms of technical content, scientific rigor, novelty, originality, and quality of presentation by at least two reviewers. Through those reviews, 18 papers were selected for publication candidates of IJIS Journal, and they were further reviewed as a Journal paper. We have three categories of IJIS papers, Regular papers, Industrial papers, and Invited papers, each of which was reviewed from different points of view. This volume includes papers among those accepted papers, which have been improved through the workshop discussion and the reviewers' comments.

We publish the journal in print as well as in an electronic form over the Internet. We hope that the issue would be of interest to many researchers as well as engineers and practitioners over the world.

Yoshiaki Terashima has been a Professor in the Department of Information Systems Science, Faculty of Science and Engineering at Soka University since 2015. He joined the Information Technology R&D Center of Mitsubishi Electric Corporation in 1984. He received his Ph.D. degree in Engineering from Shizuoka University in 2006. His research interests include distributed processing architecture, conformance testing methodologies, and network behavior estimation algorithms for IoT (Internet of Things) systems. He is a member of IPSJ, IEICE, and IEEE.

Regular Paper

Air Flow Channel Planning for Droplet-Spray-Type Olfactory Displays Using a Small Wind Tunnel

Yohei Seta[*,***], Mitsunori Makino[**], Yuichi Bannai[***], and Motofumi Hattori[***]

[*]Graduate School of Science and Engineering, Chuo University, Japan
[**]Faculty of Science and Engineering, Chuo University, Japan
[***]Department of Information Media, Kanagawa Institute of Technology, Japan
{a15.s845, makino.fme7}@g.chuo-u.ac.jp
{bannai, hattori}@ic.kanagawa-it.ac.jp

Abstract - To develop an olfactory display with high fragrance presentation performance and ease of maintenance for long-term use, a small wind tunnel was constructed to observe the inside of the flow channel of a prototype with a piezoelectric droplet spray mechanism for solving the problem of poor volatilization of the fragrance solution in the flow channel. The cross-sectional dimensions of the flow channel minimize the time for visible droplets to exist in the channel after purified water is sprayed in a small wind tunnel, where the channel height and width can be changed.

Spraying was performed on the obtained cross-sectional dimensions in five fragrance solutions, and no volatilization defects were observed in all but one of them. The piezoelectric olfactory display with the cross-sectional dimensions of the flow path obtained in this study is expected to contribute to the development of olfactory displays that can precisely synchronize the aroma with video and audio information.

Keywords: olfactory display, pulse ejection, droplet atomization, visualization

1 INTRODUCTION

Among the devices that present stimuli to people's five senses, a device that presents visual stimuli is called a display, whereas a device that presents olfactory stimuli is called an Olfactory Display. In 1962, the Sensorama [1] was the first known example of an olfactory display. Since the 1990s, multimodal interfaces have been actively researched in the field of human interfaces, and interest in the five senses of people has increased [2]. Various types of olfactory displays have been developed since the 2000s [3]-[4], and their applications include a system that presents smells synchronized with TV images and movies, with smells shown in movie theaters [5]. In virtual reality, by presenting sensory stimuli in addition to visual and auditory stimuli, users can experience smells in a VR environment [6].

Olfactory displays can be broadly classified into two types: those that control the injection of air mixed with fragrance components in advance and those that atomize solutions containing fragrance components as tiny droplets and vaporize them in air at room temperature. In either format, the user inhales air-containing fragrance components to allow them to reach the olfactory cells in the nasal cavity, thereby enabling the presentation of fragrance information [7].

When presenting an odor synchronized with video or VR content, if a high concentration of odor is continuously given that everyone can recognize, the odor will remain in the space, which is a residual problem [8]. If another odor is present in a residual state, the odor will be mixed up. Another problem is olfactory adaptation, which is the inability to perceive odors when the odor stimuli are administered on a sustained basis.

To solve these problems, Okada et al. developed an inkjet olfactory display [9] that used a thermal inkjet printer head to spray fragrance droplets. This is characterized by the fact that several hundred microscopic holes in the printer head eject several picoliters of droplets at a frequency of several KHz, enabling accurate control of the ejection volume and density over a wide dynamic range.

To confirm the basic performance of this olfactory display, experiments were conducted on the sensory characteristics of smell through subjective evaluation, and the following findings were obtained: Kadowaki et al. showed that humans could adequately perceive fragrances when pulsed fragrances were presented by injection for a very short period (approximately 0.1 s). They noted that this allowed the experiment to be conducted with a very small amount of aroma injection, thereby reducing the residual aroma in the air [10]. In addition, although the human sense of smell adapts to aroma stimuli in a short period of time, it is possible to reduce the onset of olfactory adaptation by presenting minute amounts of aroma in a pulsed (intermittent) form [11]. Noguchi et al. showed that there is a time threshold for the fragrance injection interval at which two pulsed fragrances can be perceived separately during one breath (two-point fragrance separation threshold) [12]. Based on the results of olfactory sensitivity measurements over time within one breath, they proposed an injection method that maximizes the perceived fragrance intensity during one breath. Aruga et al. proposed a method to measure scent discrimination thresholds using a pairwise comparison of two scent pulses as a measure of olfactory sensitivity, and applied it to an experiment to investigate the effect of self-motion sensation on olfactory sensitivity [13].

These findings indicate that pulse injection of scents is more effective than continuous presentation of scents in synchronization with other information media [9]. Thus, the inkjet olfactory display is superior in terms of granularity and precision of injection volume and injection time, and can be said to have a sufficient level of performance for olfactory displays. Although ink cartridges are commercially available, they are black boxes and cannot be disassembled and cleaned to deal with head clogging and odors over long periods of use, which is a maintenance problem [8]. This problem may cause inkjet olfactory displays that have been used for a long period of time to receive extra fragrance components, even when fragrances are not present, leading to a decline in fragrance switching performance and experimental accuracy. In addition, in thermal inkjet printers, the heat from the heater creates bubbles in the liquid, which push out droplets through pressure changes, and there are concerns regarding the effect on heat-sensitive fragrances. Printer manufacturers hold patents for inkjet technology, including technologies for preventing head clogging and removing clogging, and there are significant obstacles to the development of olfactory displays using new inkjet engines in the future.

Therefore, for the practical application of olfactory displays based on open technology, this research group is conducting research and development of a new olfactory display focusing on spraying droplets using the ultrasonic vibration of a piezoelectric element with microscopic holes [14], aiming at precise droplet spraying based on a simpler principle. In this study, this is referred to as a piezoelectric olfactory display. The piezoelectric olfactory display is intended to be applied not only to VR content but also to a wide range of fields, such as psychophysical experiments, by achieving both a pulse injection performance of fragrance equivalent to that of the inkjet type and maintainability that allows long-term use.

In this process, a prototype was made following the body channel dimensions used for the inkjet-type [15]. When the prototype was sprayed with droplets for a short period to evaluate its performance, droplet adhesion on the bottom surface of the channel was confirmed [16]. The droplets were confirmed to vaporize afterwards, but some droplets remained for a few seconds after the end of droplet spraying. This indicates the possibility of continuous scent presentation for more than ten times the time of droplet atomization. This problem must be solved to achieve a pulse injection performance of a fragrance equivalent to that of an inkjet-type.

Possible solutions to this problem include changing to a more volatile fragrance solution, reducing the diameter of the sprayed droplets to improve their surface area, increasing the volatility, and expanding the flow path shape such that the sprayed droplets do not contact each other until they fully evaporate. In the case of a change in the fragrance of a solution, a change in its physical properties may interfere with the atomization of the piezoelectric element. In addition, to reduce droplet diameter, it is necessary to perforate the piezoelectric element mold with micropores of different sizes. When changing the shape of the flow path, it is necessary to find a flow path dimension that is sufficiently volatile, even if the atomizing object comes into contact with it; however, this will lead to a larger body.

In this study, we employed a change in channel geometry, which has the lowest change cost among the above approaches, to obtain the necessary and sufficient channel cross-sectional dimensions without droplet adhesion during droplet spraying. An experimental wind tunnel with variable channel walls was constructed to allow direct observation of droplet volatilization failure in the channel. In the experimental wind tunnel, the structure was created using fluid dynamics to achieve a stable and unbiased airflow inside the tunnel [17]-[18]. Based on the hypothesis that if there is no droplet adhesion in purified water, there will be no droplet adhesion in fragrance solutions containing a portion of ethanol. We will conduct a purified water spray experiment while moving along the channel wall and confirm whether there is droplet adhesion by visually observing and extracting the sprayed object from the video captured using image processing technology. When the flow dimensions without droplet adhesion of water were obtained, the effectiveness was verified by injecting a fragrance solution with the same dimensions. By creating a piezoelectric olfactory display with the obtained channel dimensions, we expect to make progress in the development of an olfactory display with a droplet spray mechanism that is easy to maintain and has the same pulse injection performance as an inkjet-type fragrance solution and easy removal of fragrance.

2 METHODOLOGY

2.1 Piezoelectric Olfactory Display

Our research group developed a piezoelectric olfactory display that adopts a piezoelectric element as a droplet atomization mechanism [14] to circumvent the disadvantages of inkjet olfactory displays such as thermal injection, difficulty in downsizing the atomization mechanism, and complicated cleaning and maintenance [9].

Figure 1 illustrates the operation of the piezoelectric olfactory display [15]. This device uses a fan to blow a liquid fragrance that is atomized by a piezoelectric device built into a liquid fragrance tank to present the scent to the user. The piezoelectric device had a diameter of approximately 1 cm, and 100 micropores of 9 μm diameter were drilled in its center. Although the liquid fragrance in the liquid fragrance tank does not pass through the micropores owing to surface tension when no voltage is applied, it is atomized through the micropores by vibrations when a voltage is applied. The piezo element oscillated once every 10 μs, and approximately 167 pL of pure water atomization was observed with each oscillation.

A fan was employed to blow air over the atomized liquid fragrance, promote its volatilization of the liquid fragrance, and present the scent to the user.

A 3 mL capacity liquid fragrance tank was placed above the piezoelectric device and the fragrance in the tank was continuously ejected from the piezoelectric device. The tank, piezoelectric device, and the piezoelectric device holder at the bottom of the tank were removable to facilitate cleaning and maintenance. An experiment to measure the mass of purified water atomized 10 times for 10 s verified that the average injection mass was 0.17 g (standard deviation: 0.0077) and remained constant regardless of the number of atomization. Different types of fragrances have different viscosities,

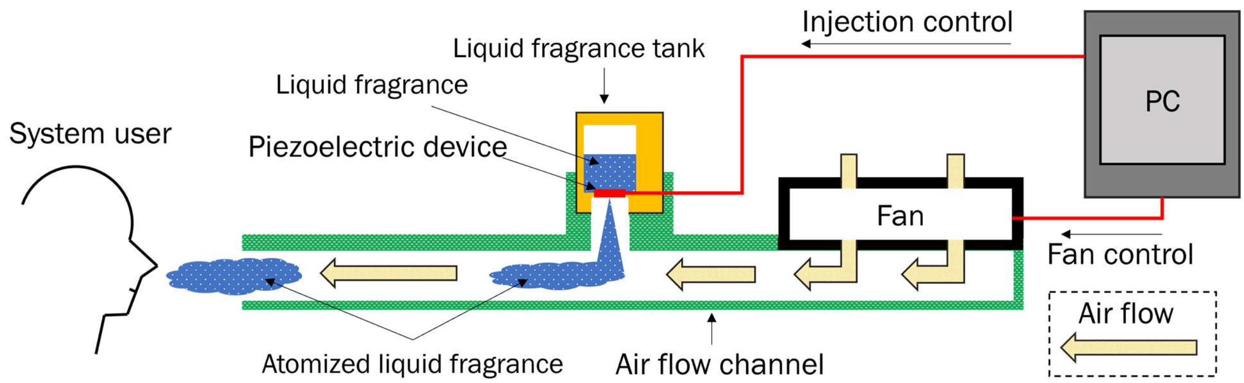


Figure 1: Cross-sectional image of piezoelectric olfactory display (side).

surface tensions, and other characteristics; therefore, fragrance solutions to be atomized in olfactory displays must be adjusted according to their composition.

2.2 Limitations of the Piezoelectric Device Olfactory Display

Our prototype's air flow channel has a cross section 20 mm high and 60 mm wide, which follows the structure of the inkjet olfactory display [9], [15]. To install the liquid fragrance tank, a 10 mm rise section was provided at the top of the air flow channel, and a cylindrical hole with a diameter and height of 15 mm and 15 mm, respectively, was created in the rise. The surface of the liquid fragrance tank was attached to the top of the raised section and atomization was directed vertically toward the bottom of the airflow channel. To prevent the piezoelectric device from being directly touched by hand, a 5 mm high dimple exists at the bottom of the liquid fragrance tank as a frame to protect the surface of the piezoelectric device, and a space for fixing the piezoelectric device exists above the frame. Therefore, the distance from the atomization position of the piezoelectric device to the bottom of the channel is 40 mm, which is equivalent to the atomized height of an inkjet olfactory display, and previous experiments with inkjet in olfactory displays, atomized substances are sufficiently vaporized in airflow channels [10]. Therefore, it was assumed that the atomized substance used in the piezoelectric olfactory display experiment vaporized immediately after ejection and did not remain as droplets on the wall surface.

However, the atomization mechanisms of inkjet and piezoelectric olfactory displays differ, including the composition of the liquid fragrance solution; hence, the vaporization of the atomized substance is also expected to differ.

We created an acrylic plate airflow channel with a height of 20 mm and a width of 60 mm to confirm the conditions in the airflow channel immediately after the atomization of droplets in the piezoelectric device olfactory display of our prototype [16]. We connected the liquid fragrance tank of our prototype method to the acrylic plate airflow channel and observed the inside of the airflow channel by capturing video images at the time of atomization. In the 0.3 s pulse injection experiment, purified water clearly left droplets over a wide area at the bottom of the channel, and it took approximately 10 s before the droplets were no longer visible. The droplets of the banana fragrance solution were not visible earlier than those of the purified water, with an average of 4.8 s. The time required for the droplets of the banana fragrance solution to disappear was approximately 10 s, whereas that of purified water was approximately 10 s. The results showed that the atomized substance adhered to the bottom surface of the conventional structure as droplets at the time of atomization of the banana liquid fragrance and purified water and evaporated later than after the system had finished atomizing them.

To accurately determine a person’s scent detection threshold using a droplet-atomizing olfactory display, it is necessary to switch between odorless and attached states with a high degree of accuracy [9]. If the presentation state by the system is odorless but the subject is presented with a scent because of the above problem, the olfactory characteristics of the subject may not be measured correctly.

3 METHOD

To improve the volatilization of droplets observed in the piezoelectric olfactory display prototype, an experimental apparatus that allowed visual inspection of the inside of the display was constructed to confirm the conditions at the time of spraying. Based on the hypothesis that purified water is less likely to volatilize than fragrance solutions based on the residual time of droplets observed in the prototype described in Chapter 2, we created an experimental apparatus with a fixed channel width and changeable channel height in Experiment 1 to find a channel height at which droplet adhesion of purified water on the bottom surface is not observed. The channel height was determined to prevent droplets of purified water from adhering to the channel bottom. Because droplet adhesion was observed in the width direction in Experiment 1, Experiment 2 was conducted to determine a channel width that could be changed while keeping the channel height fixed at the height obtained in Experiment 1, so that no droplet adhesion could be observed on the sides. The presence or absence of droplets is determined by image processing, which highlights the droplets when visual observation is impossible. The effectiveness of the method was verified by spraying five fragrance solutions of different compositions on the final channel dimensions.

3.1 Experiment 1: Confirmation of the Required Air Flow Channel Bottom Distance During Purified Water Atomization

An experimental airflow channel was constructed using acrylic plates to determine the minimum distance at which a droplet was atomized by the piezoelectric device without contact with the bottom of the airflow channel. Figure 2 shows the configuration and dimensions of the experimental airflow channel.

In the experiment, the distance from the piezoelectric device to the bottom of the airflow channel was varied, and a video of the droplet atomization was recorded. Video images of the atomized droplets were captured and checked for the presence or absence of droplet adhesion at the bottom of the airflow channel. The threshold value of the bottom height for the presence or absence of deposits at the bottom of the airflow channel was defined as the minimum distance without contact with the bottom of the airflow channel when atomizing the droplets.

Previous studies have demonstrated that liquid fragrances containing ethanol or other solvents vaporize faster than purified water [16]. By determining the threshold of the bottom height relative to the purified water, liquid fragrances that could be used in the future can be considered. Therefore, only purified water was used in this experiment.

The experimental airflow channel employed a bottom surface whose height could be modified as desired; this is referred to as a movable floor in this study. The movable floor was an acrylic plate with the same dimensions as the bottom of the airflow channel, wrapped with vinyl tape at both ends and fixed in position by friction with the acrylic plate on the channel side. Therefore, the movable floor could be easily removed by changing the bottom height. To maintain constant conditions on the movable floor surface during continuous injection, each time the bottom height was changed, the surface was wiped with a paper rag moistened with purified water, and then dried again with a dry paper rag. While fixing the movable floor, the height of the installed surface was checked using a digital caliper, and a level was used to check the level at the part protruding from the channel prior to the experiment.

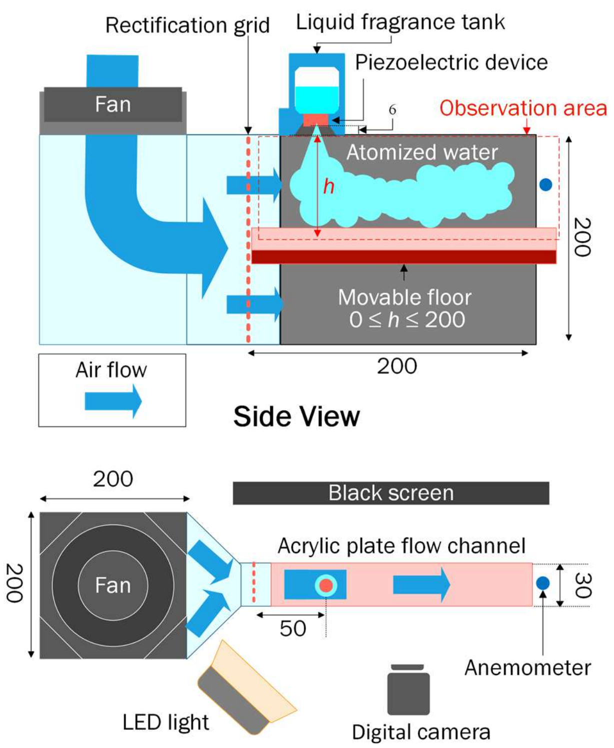


Figure 2: Structure of experimental channel for confirmation of the required airflow channel height (unit: mm).

The air velocity in the airflow channel during the droplet atomization experiment was 1.8 m/s, which was adopted in conventional experiments on inkjet olfactory displays [10].

A PWM-controllable PC fan was utilized to provide a constant airflow in the experimental airflow channel. The selection of the fan was based on the maximum airflow rate, and the fan was required to sufficiently satisfy the target air velocity over the entire experimental airflow channel cross section. Blowing air using rotating fans generated low velocity near the center of the air velocity distribution, asymmetrical bias, and turbulence. The airflow in the experimental airflow channel was stabilized to improve the reproducibility of the experimental results.

In this experimental airflow channel, a grid-like structure [17]-[18] used for rectification in the experimental wind tunnel was installed in the cross-section of the airflow channel, which flowed into the airflow channel from the fan mounting area. Previous experiments have demonstrated that the grid shape is effective in stabilizing the airflow in the airflow channels of piezoelectric olfactory displays when the pitch width, plate thickness, and open area ratio are 5 mm, 1 mm, and 60%, respectively [19]; in addition, the same parameters will be adopted in this study. The fan was mounted in the direction of the intake from the top of the experimental apparatus and perpendicular to the cross-section of the airflow channel to reduce the influence of the low-speed portion at the center of the fan. During the experiment, an anemometer was placed at the center of the cross-section, which was at the end of the experimental airflow channel, and the power of the fan was adjusted to achieve the target wind velocity while checking the wind velocity. The experiment was conducted only at wind speeds of 1.8 ± 0.03 m/s.

3.1.1 Experiment 1: Pre-Confirm of Wind Velocity Distribution in the Experimental Air Flow Channel

To verify the maximum air velocity and rectification effect in the experimental airflow channel, and to check whether the use of the movable floor significantly disturbed the air velocity distribution in the airflow channel cross-section, the air velocity distribution was measured at the vertical and horizontal centers of the cross-section at the airflow channel end when the experimental airflow channel height was h = 200 mm and h = 100 mm, with the movable floor at the output under the maximum airflow rate of the fan.

The experimental channel was created, and a 180 mm PC case fan “AP184i PRO” from Silver Stone was used for the fan, with PWM control using an Arduino uno 3. The experiment was conducted at a voltage of 12 V, which was

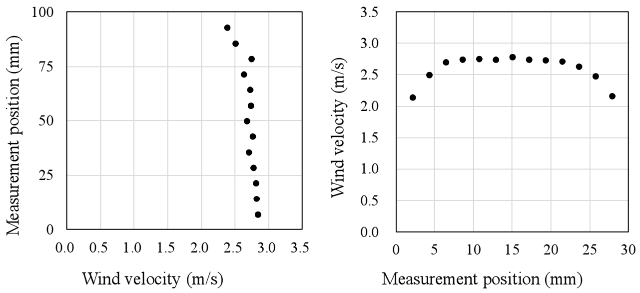


Figure 3: Wind speed distribution (h = 100 mm).

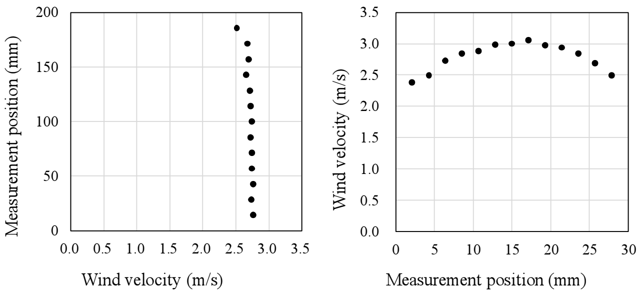


Figure 4: Wind speed distribution (h = 200 mm).

Table 1: Average wind velocity of channel cross-section

	Horizontal	Vertical	Horizontal	Vertical
h (mm)	100		200	
Mean (m/s)	2.61	2.7	2.8	2.7
SD (m/s)	0.21	0.12	0.21	0.06

the rated output of the fan, to determine the average wind speed at which a stable output was possible. The anemometer used was a KANOMAX Thermal Anemometer (Anemo Master Light Model 6006). The center of the anemometer element was captured for 10 s at each sampling point of the flow cross section, and the average

The value recorded per second was the measured value at each point. The anemometer element has a length of approximately 3 mm. During the measurement, the wind anemometer was oriented perpendicular to the direction in which the sampling point moved along the longitudinal axis such that the wind speed values were recorded as points.

In both cases, the vertical wind velocity distribution was almost flat, with an average of approximately 2.7 m/s, the horizontal wind velocity distribution was almost symmetrical, with the fastest near the center and gradually slowing as it approached the wall (Figs. 3-4). Table 1 presents the average wind velocity and standard deviation of the cross-section of the experimental airflow channel.

The obtained results indicate that the experimental airflow channel can be utilized for this experiment because the target air velocity of 1.8 m/s is sufficiently satisfied, and there is negligible turbulence in the air velocity distribution owing to the asymmetric flow caused by fan rotation and the influence of the movable floor.

3.2 Experiment 2: Confirmation of the Required Air Flow Channel Width During Purified Water Atomization

From the results of Experiment 1, no adhesion of purified water to the bottom of the channel was observed when the distance from the bottom of the channel to the piezoelectric element was at least 70 mm. In addition, droplet adhesion to the channel surface occurred regardless of the channel height when the width of the experimental channel was 30 mm. Therefore, to obtain a channel shape that does not allow droplet adhesion, this study describes a new experimental channel whose side width can be changed to obtain a channel width that does not allow droplet adhesion.

Based on the results of Experiment 1, the experimental channel shape was fixed at a channel height of 70 mm, and the total distance from the piezoelectric element to the bottom of the channel was 76 mm, including a 6 mm depression in the liquid fragrance tank. The channel width can be varied within the range 30-70 mm. Because the flow beyond a channel length of 200 mm could not be observed at the Experiment 1, a channel length of 300 mm was used in Experiment 2.

Using the PC fan connection module used in Experiment 1, we created a new acrylic plate flow channel and a rectifying flow structure. Two acrylic panels were placed parallel to each other in the channel and the distance between the panels was manually adjusted from the outside to change the channel width, which was used as a movable sidewall. When a prototype of the movable sidewall was fabricated using a 2-mm-thick acrylic plate, it was confirmed that the acrylic plate was warped, making it difficult to install it parallel to the wall. Because warpage was reduced using a 3-mm-thick acrylic plate, a 3-mm-thick acrylic plate was used for the movable sidewall in this method.

In a channel using only a rectifying grid with an aperture-to-area ratio of 60% and a grid spacing of 5 mm as the rectifying structure, the distribution was biased in the horizontal direction. This biased distribution can be attributed to the sudden change in the cross-sectional area of the flow path from the cross-section of a 200 mm high, 200 mm wide acrylic box connected to a fan to an experimental flow channel with a height of 70 mm and width of 80 mm. To weaken this effect, a bell-mouth shape was adopted, which enabled stable fluid intake by smoothly decreasing the cross-sectional area of the flow channel while connecting it from a wide space to a narrow duct. To prevent sudden contraction of the cross-sectional area of the experimental flow channel after it is narrowed by the movable sidewall at the connection between the bell mouth and the experimental flow channel, a hole is provided where the handle of the movable sidewall overlaps the cross-sectional area of the flow channel to allow air to pass through.

In this experimental channel, the wind velocity distributions at w = 70 mm and w = 35 mm were analyzed in the same way as described in Experiment 1 Both of mean wind speeds in horizontal directions were 2.51 m/s, the maximum wind speed differences at the horizontally symmetrical positions were 0.18 m/s and 0.14 m/s, and the standard deviations were 0.08 m/s and 0.17 m/s. The mean wind speeds in vertical

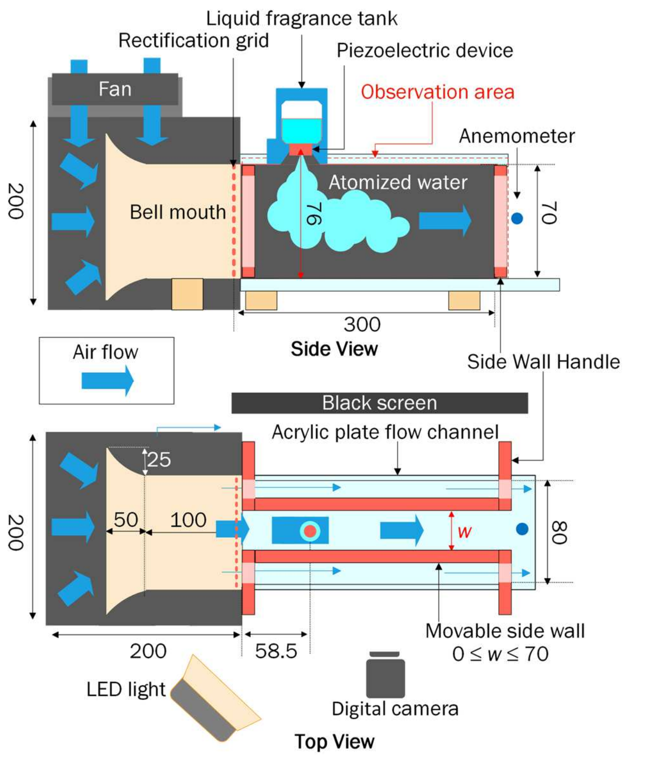


Figure 5: Structure of experimental channel for confirmation of the required air flow channel width (unit: mm).

directions were 2.57 m/s and 2.65 m/s, and the maximum wind speed difference at the vertical target position was 0.32 m/s and 0.30 m/s, with standard deviations of 0.08 m/s and 0.09 m/s. These wind velocity measurements did not reveal any extreme fluctuations in the wind velocity distribution due to changes in the channel width. The maximum wind velocity difference at the symmetrical horizontal position was 0.33 m/s at a channel width of 70 mm before the bell mouth was installed, whereas it was approximately half that after the bell mouth was installed. The wind velocity bias of each channel was considered acceptable and the experimental channels were used for the experiment.

Figure 5 shows an image of the experimental apparatus. Using this experimental channel, atomization experiments of pure water were conducted at 5 mm intervals at a channel width ranging from 30 mm to 70 mm. When the width of the flow channel was changed, the movable sidewalls were operated. Moreover, the width of the channel was measured with a digital caliper to confirm that the movable sidewalls were parallel to each other and that the experimental channel was horizontal using a leveler before conducting the experiment. In each channel-width experiment, the droplets were first atomized for more than 10 s and observed with the naked eye at a close distance of approximately 100 mm from the channel wall to check for abnormalities in the experimental equipment, droplet atomization, and droplet adhesion on the channel wall.

After confirmation, four sets of videos were captured from the side surface of the channel, with one set consisting of three consecutive droplet atomization for 1 s at 5 s intervals. An anemometer was used to check the wind speed at the center of the cross-section at the channel tip. Similar to the setup in Experiment 1, the camera was ready to start capturing the wind speed when it was 1.8 ± 0.03 m/s. Each set used different camera settings. In this experiment, full HD resolution (240 fps), 4 K resolution high dynamic range setting (30 fps), 4 K resolution normal capture (60 fps) brightness adjustment based on bright points, and 4 K resolution normal capture (60 fps) brightness adjustment based on dark points were used. After the experiment, the following were confirmed based on the results for each channel width: No obvious droplet adhesion was observed by naked-eye inspection. In each of the 12 videos of droplet atomization, the target frame was captured 0.17 s after the end of atomization when an object moving at 1.8 m/s had traveled the full distance of 300 mm from the end of ejection.

When a channel width that simultaneously satisfied these criteria was obtained in the atomization experiment with purified water, the atomization experiments of solutions with five different water-soluble fragrances (banana, lemon, rose, vanilla, and blueberry) were conducted at the minimum channel width in the same manner as the method used for purified water.

The atomizing solution for these experiments was prepared using commercially available water-soluble flavoring agents for food additives from Marugo Corporation, as described in our prototype. For the fragrance solution, dilute the water-soluble fragrance of each flavor with purified water, obtain a fragrance component to be 5% by mass. The substances and amounts of the flavor components of each flavor used in the experiment are not disclosed. Therefore, based on the mass of the indicated organic solvent, the difference between the overall mass and the mass of the organic solvent was assumed to be the mass of the flavor components, and the ratio of the fragrance solutions mixed with purified water for each flavor was determined. Table 2 lists the ratios of the solutions used in the experiments. Of the five fragrance solutions selected for the experiment, two (banana and lemon) have been used in a previous study. Among the newly prepared fragrance solutions, we selected three (rose, vanilla, and blueberry) that differed significantly from those used in the previous study in terms of the physical properties (surface tension and viscosity) and composition of the fragrance solutions. If this experiment shows similar to those obtained for purified water, the experimental channel width and channel height can be considered appropriate shape parameters for the design of a rectangular channel to be used in novel piezoelectric olfactory displays with a central air velocity of 1.8 m/s at the tip channel cross-section.

Table 2: Component ratios of atomization solutions (%).

	Vanilla	Rose	Banana	Blueberry	Lemon
Fragrance Components	5.00	5.00	5.00	5.00	5.00
Ethanol	2.24	4.00	5.11	5.00	7.92
Glycerol	0.22	1.00	1.14	-	0.97
Propylene Glycol	-	-	0.11	2.50	-
Water	92.54	90.00	88.64	87.50	86.11

3.3 Image Processing for Improved Visibility of Droplets

To confirm the presence or absence of droplet adhesion on the atomized target and channel wall, a background subtraction method was used to visualize the atomized target, using the frame before the start of droplet atomization as a background image. The contrast-enhanced image was then observed using histogram flattening to confirm the absence of droplet atomization-caused droplet adhesion to the channel. For image processing, Python3 and Open CV 4.6.0 library were used.

The "cvtColor method" with the COLOR_BGR2GRAY parameter was used to convert the target frame and background images to grayscale and the RGB values of the pixels to luminance values. The difference between the luminance values of the gray scaled target frame and the background image is obtained using the "absdiff method." This method only visualizes areas in which the luminance values change between the target and background images.

Histogram averaging was performed on the images obtained using the background subtraction method to enhance the contrast and improve the visibility of the observed object. The method uses the "clahe method" to perform adaptive histogram averaging on an image obtained by subtracting the background image from the target frame.

4 RESULTS

4.1 Experiment 1: Confirmation of the Required Air Flow Channel Bottom Distance During Purified Water Atomization

An experimental airflow channel was fabricated using a 3D printer and an acrylic sheet.

Figure 6 shows the atomization at h = 34 mm. A digital camera was positioned at the center of the screen at an angle that captured the entire movable floor surface, and any droplet adhesion on the bottom of the channel during purified water spraying was illuminated by white LED light adopted as illumination and visualized in white.

Figures 7-9 show present the experimental results for h = 34, 63, and 64 mm, respectively, during 1 s of purified water

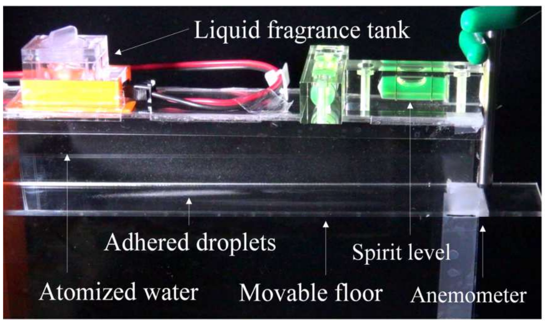


Figure 6: Illustration of experimental results (h = 34 mm).

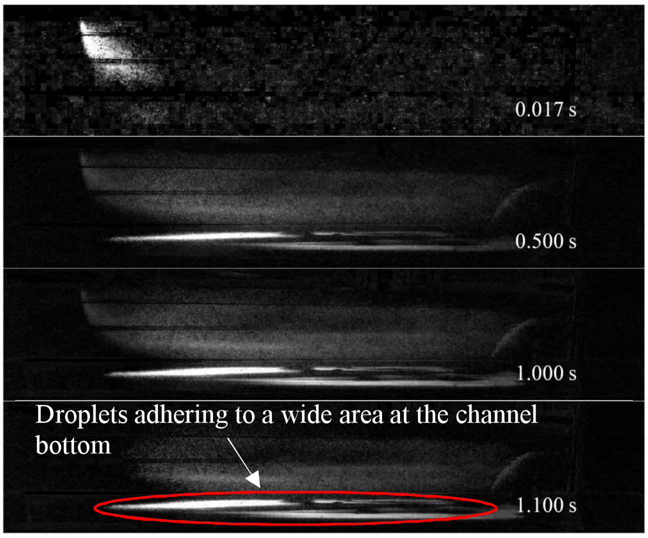


Figure 7: Experimental results (h = 34 mm, Color difference from 0 s and after contrast adjustment, Clipping only in air flow channel).

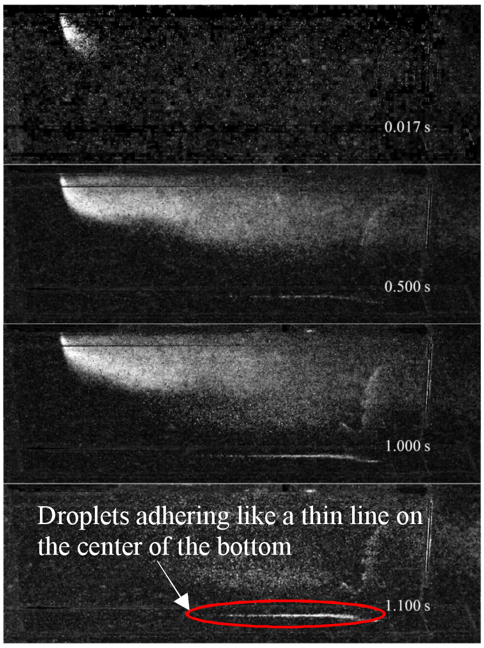


Figure 8: Experimental results (h = 63 mm, Color difference from 0 s and after contrast adjustment, Clipping only in air flow channel).

atomization, where; the time series changed when the atomization start frame was set to 0 s.

These images were clipped only inside the airflow channel, and the difference in color values between the atomizing start frame and each frame was determined using the method of section 3.3. The color values were output only for the areas that changed from the initial frame; and the areas that did not change were output in black. At h = 34 mm (in Fig. 7), droplet adhesion was clearly observed over a wide area of the movable floor surface. The droplets remained at the bottom of the air flow channel for 1.1 s after atomization was completed, and the atomized substance flowed out of the airflow channel. As shown in Fig. 8, at h = 63 mm, linear droplet adhesion was observed at the center of the movable

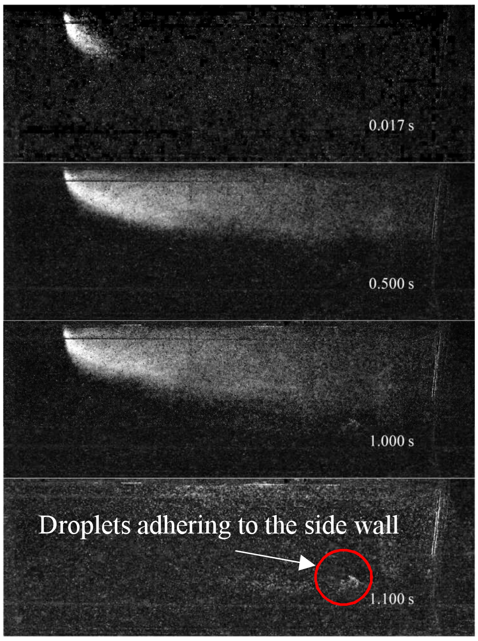


Figure 9: Experimental results (h = 64 mm, Color difference from 0 s and after contrast adjustment, Clipping only in air flow channel).

floor in front of the airflow channel in the form of a thin line. In Fig. 9, h = 64 mm, the atomized substance flowing in the channel was confirmed; however, no bottom adhesion was observed, and the same result was obtained in six subsequent trials conducted for confirmation.

In all cases, however, droplet adhesion was observed on a surface on the channel side 1.1 s after the end of injection, when the 1.8 m/s wind velocity should have sufficiently ejected substances out of the channel.

This means that the 30 mm channel width of the experimental channel was insufficient to prevent droplet adhesion to the channel sides.

4.2 Experiment 2: Confirmation of the Required Air Flow Channel Width During Purified Water Atomization

An experimental channel with a variable width was created, and purified water atomization experiments were conducted. Similar to the setup of Experiment 1, a shape model of the experimental channel was created using Autodesk's "Fusion360" software. A bell mouth was made from the shape model using a 3D printer ("Raise 3D E2") with a PLA filament as the material. The experimental flow channel, rectifying grating, and movable sidewalls were fabricated by cutting acrylic sheets using a laser cutting machine based on the shape data. The parts were fixed with an acrylic adhesive and sealed with vinyl tape to prevent air leakage from the joints.

In previous experiments, droplet atomization in liquid fragrance tanks was found to fluctuate owing to the deterioration of the piezoelectric element, the presence of dirt, and poor contact between the connection terminals. The droplet-atomizing performances of the four liquid fragrance tanks were compared. The liquid fragrance tank with the

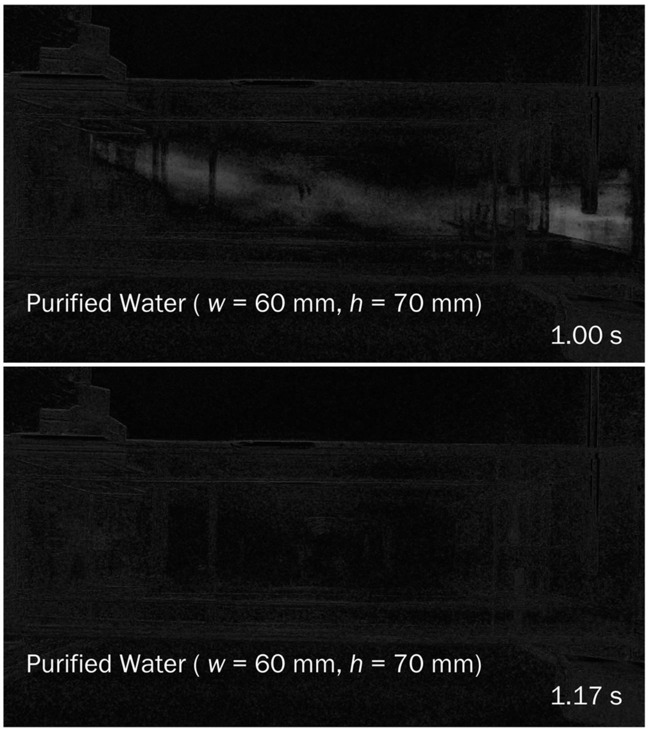


Figure 10: Side view of experimental results of purified water atomization (images obtained by Contrast Limited Adaptive Histogram Equalization for background subtraction before and after atomization).

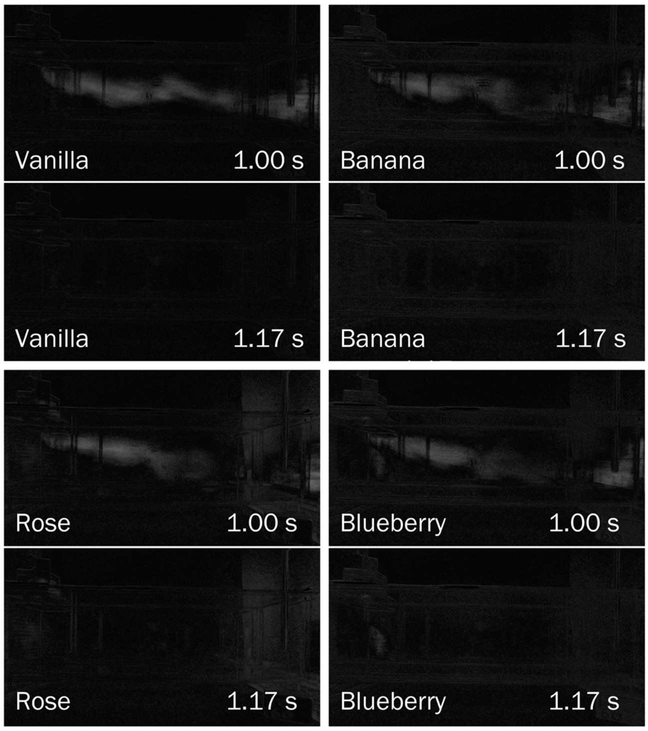


Figure 11: Side view of experimental results of four fragrance solution atomization (w = 60 mm, h = 70 mm).

highest average atomizing mass and the lowest coefficient of variation was selected for use in the experiment. The average atomizing mass of purified water droplets in the liquid

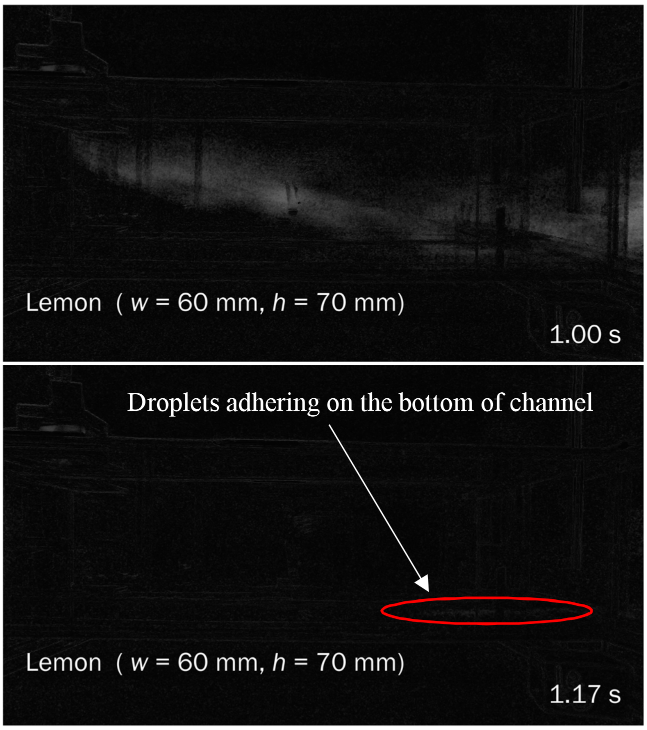


Figure 12: Side view of experimental results of the lemon fragrance solution atomization.

fragrance tank used in the experiment was 2.08 mg/s, with a standard deviation of 0.08 mg/s and a coefficient of variation of 3.60%. The following results were obtained using one of the liquid fragrance tanks. A piezoelectric element and a liquid fragrance tank containing purified water were placed in the experimental channel such that the piezoelectric element surface was horizontal and fixed with vinyl tape to prevent displacement during the experiment.

During the filming, the lights were turned off, the area around the experimental apparatus was covered with a blackout curtain to prevent the influence of flickering natural or fluorescent light, and a small LED light source was used for video recording. Purified water atomization experiments were conducted using the experimental setup described above. A DJI "OSMO ACTION" camera was used to capture images of the experimental flow channel while switching the camera settings for each experiment. The temperature and humidity at the start of the experiment were 21.6°C and 20.7%, respectively.

The results of the purified water atomization experiments are presented below. Naked-eye inspection of the channel wall during 10 s of purified water atomization revealed numerous small droplets on the channel wall and top surface at the center of the channel at *w* = 30-35 mm. The small droplets that adhered to the wall volatilized immediately after contact. At *w* = 40-55 mm, droplet adhesion to the top surface of the channel was no longer observed. Small droplets were observed to contact the channel-side surface in the range of 100-150 mm from just below the liquid fragrance tank toward the channel tip and volatilized immediately after contact. The frequency of the droplet contacts with the channel-side surface decreased as the channel width increased. At *w* = 60-70 mm, almost no droplet contact was observed by naked-eye inspection. The target frame was captured 0.17 s after droplet atomization by applying image processing to a video captured at *w* = 60 mm. Figure 10 shows the results after image processing just before the end of 1 s of purified water atomization and 0.17 s after the atomization.

Next, experiments were conducted at *w* = 60 mm by atomizing the solutions of five water-soluble fragrances (banana, lemon, rose, vanilla, and blueberry).

No droplet adhesion to the channel sidewall was observed for any of the liquid fragrance solutions, either with the naked eye or in the target frames of the video after image processing. However, droplet adhesion to the bottom of the channel was observed in the 11 atomization of the lemon fragrance solution. White streaky adherent droplets were observed in the center of the bottom of the channel from 150 mm to 200 mm from the atomizing start position to the end of the channel after ≥ 0.17 s. Figures 11-12 show the results after image processing just before the end of 1.00 s of liquid fragrance solutions atomization and 0.17 s after the atomization.

5 ANALYSIS

5.1 Experiment 1: Confirmation of the Required Air Flow Channel Height at the Time of Purified Water Atomization

The airflow channel height of the conventional piezoelectric olfactory display (*h* = 34 mm) was insufficient because of the large number of droplets adhering to the bottom surface. Bottom adhesions were observed at $h \leq 63$ mm; however, no bottom adhesions were observed at $h > 64$ mm even after repeated atomization. The required air flow channel bottom distance at the time of purified water atomization at a flow velocity of 1.8 m/s, when vibrating at 100 kHz using a piezoelectric device with 100 micro-holes of 9 µm, was verified to be 70 mm, including the dimple in the liquid fragrance tank.

In all cases, contact was observed on the sides during atomization at an 30 mm airflow channel width.

5.2 Experiment 2: Confirmation of the Required Air Flow Channel Width at the Time of Purified Water Atomization

Experiments using purified water and four kinds of except lemon fragrance solution showed that no droplet residue was observed on the wall surface when droplets were atomized by using a channel with geometry parameters of $w \geq 60$ mm and $h \geq 70$ mm. This confirms that for water-soluble liquid flavorings other than lemon flavorings used in the experiment, there are no visible droplets in the channel at 0.17 s after the end of spraying, due to the obtained channel shape parameters, rather than the conventional parameters of *w* = 60 mm and *h* = 34 mm. Compared to the 4.8 s required for the volatilization of banana flavor droplets observed in the prototype, the residual time of the droplets in the channel was reduced to approximately 3.5%. The ambient fragrance during all the atomization experiments using water-soluble liquid fragrances qualitatively indicated that each fragrance could be represented by the piezoelectric olfactory display.

However, the lemon fragrance solution used in this experiment in a channel with a flow channel more than 200 mm from the atomizing point may cause adhesion problems at the bottom of the channel.

In this experiment, we considered why bottom adhesion occurred only with the lemon fragrance solution. Considering the volatility of the liquid from the boiling points under standard atmospheric pressure of the components of the lemon fragrance solution, glycerin (boiling point 290°C), which has the highest boiling point among the known components, was present at approximately 1%. However, under the same conditions, no adhesion to the bottom surface was observed for rose and banana fragrance solutions containing similar amounts of glycerin. The lemon fragrance solution contained the highest amount of ethanol (boiling point 78.37°C) and the lowest amount of water (boiling point 100°C) among the tested solutions. The unique flavor components of the lemons used in this study are not disclosed. Therefore, it is difficult to explain the reason for droplet adhesion to the bottom surface because of the volatility of the ingredients. In the future, it will be necessary to conduct experiments using lemon fragrance solutions with known ingredients and to measure the actual volatilization time of the fragrance solutions.

The effect of specific gravity of the fragrance solution was also considered. The different specific gravities of the droplets may cause differences in the diffusion range owing to the effect of gravity. All video frames for a certain period of time during atomization were superimposed and added together to visualize the diffusion range of the atomized droplets. This may explain the difference in the diffusion range of each fragrance solution, and the relationship between the specific gravity of the fragrance solution and the diffusion range of the droplet, which may explain the lower adhesion of the lemon fragrance.

In the movie taken from the experiment using purified water with a channel width of w = 30-55 mm, where naked-eye inspection confirmed droplet adhesion, droplet adhesion was not as clearly identified as in Experiment 1. First, there was a difference in the attachment method of the liquid fragrance tank between Experiment 2 and Experiment 1. In Experiment 2, the liquid fragrance tank was fixed to the experimental channel using vinyl tape to maintain the direction of atomization of the liquid fragrance tank perpendicular to the bottom of the channel. Therefore, in the results of Experiment 1, where droplet adhesion was observed on the side surface of the flow channel, it is possible that a tilt existed in the atomizing direction of the liquid fragrance tank, and droplets were sprayed on one side of the channel. Second, the effect of illumination on the visualization of droplets may differ from that of the experimental channel used in Experiment 1 because one more acrylic plate passes between the channel wall with attached droplets and the video equipment owing to the movable sidewalls. It indicates the need for improved lighting and experimental filming conditions.

6 CONCLUSION

To improve the performance of the prototype piezoelectric olfactory display, experiments were conducted to determine the channel height and width at which sprayed droplets were no longer visible in the channel. To correct the bias in the wind velocity distribution identified during the process of developing the experimental channel, data from wind tunnel production were used to obtain a stable wind velocity distribution with reduced bias using a rectifying grid and upon shrinkage of the channel cross-section. Using an experimental flow whose height could be changed, we conducted a purified water atomization experiment at a central wind velocity of 1.8 m/s at the channel tip, similar to the conditions of the conventional method. The results of the naked-eye inspection and experimental filming indicated no droplet adhesion to the channel bottom when the channel height was 64 mm or when the distance from the channel bottom to the piezoelectric element was 70 mm. An experimental channel with a variable width was used to conduct the purified water atomization experiment.

No droplet adhesion was observed on the channel-side surface when the channel width was 60 mm through video and visual observations. In the experimental channel with geometric parameters of w = 60 mm and h = 70 mm, where no droplet adhesion was observed in purified water, solutions of five liquid fragrances (banana, lemon, rose, vanilla, and blueberry) were atomized, and no droplet adhesion was observed on the channel sidewall, similar to the results for purified water.

However, in the lemon-flavored solution used in this study, droplet adhesion to the bottom of the channel was observed in all results. This indicates that in an experiment to determine the appropriate channel height for an olfactory display, only atomizing purified water and checking for droplet adhesion to the bottom are not sufficient to draw a conclusion. In the future, it will be effective to use this experimental flow channel to set the fragrances to be used in advance, determine the appropriate shape sufficient to use them, and check the effects of changing the mixture components of the fragrance solution to adapt to the set channel shape.

In addition, the piezoelectric olfactory display qualitatively presented the liquid fragrances used in the experiment. These experimental results indicate that for applications involving the liquid fragrance used in the experiment, other than the lemon solution, the residual droplets on the channel wall can be reduced using a channel width of $\geq$ 60 mm and channel height of $\geq$ 70 mm when designing a new appropriate channel structure. Future experiments using new fragrances will confirm whether the shape parameters obtained in this study are applicable to a wide variety of liquid fragrances. Based on the shape parameters determined in this study, we will create a prototype of a new piezoelectric olfactory display and compare with an inkjet olfactory display using a test subject.

REFERENCES

[1] M. L. Heilig, Sensorama Simulator, U.S.Patent 3050870 (1962).

[2] T. Narumi, T. Kajinami, S. Nishizaka, T. Tanikawa, and M. Hirose, "Pseudo-gustatory display system based on cross-modal integration of vision, olfaction and gustation", 2011 IEEE Virtual Reality Conference, pp.127-130 (2011).

[3] Y. Yanagida, "Projection based olfactory display with nose tracking", VR '04: Proceedings of the IEEE Virtual Reality 2004, pp.43-50 (2004).
[4] T. Yamada, S. Yokoyama, T. Tanikawa, K. Hirota, and M. Hirose, "Wearable Olfactory Display: Using Odor in Outdoor Environment", IEEE Virtual Reality Conference (VR 2006), pp.199-206 (2006).
[5] C. J. 4dplex, 4DX SIGNATURE EFFECTS, https://www.cj4dx.com/aboutus/aboutus.php (Accessed: 20-Apr-2023).
[6] A. Mochizuki, T. Amada, S. Sawa, T. Takeda, S. Motoyashiki, K. Kohyama, M. Imura, and K. Chihara, "Fragra: a visual-olfactory VR game", ACM SIGGRAPH 2004 Sketches, p.123 (2004).
[7] Y. Yanagida, "A survey of olfactory displays: Making and delivering scents", SENSORS, 2012 IEEE, pp.1-4 (2012).
[8] Y. Bannai, "Olfactory display using piezoelectric element", in Odor sensing, analysis and its visualization and quantification, Technical Information Institute, pp.518-529 (2020). (in Japanese)
[9] S. Sugimoto, D. Noguchi, Y. Bannnai, and K. Okada, "Ink Jet Olfactory Display Enabling Instantaneous Switches of Scents", Proceedings of the 18th ACM International Conference on Multimedia, pp.301-310 (2010).
[10] A. Kadowaki, J. Sato, Y. Bannai, and K. Okada, "Measurement and Modeling of Olfactory Responses to Pulse Ejection of Odors", Journal of Japan Association on Odor Environment, Vol.39, No.1, pp.36-43 (2008).
[11] A. Kadowaki, J. Sato, Y. Bannai, and K. Okada, "Presentation Technique of Scent to Avoid Olfactory Adaptation", 17th International Conference on Artificial Reality and Telexistence (ICAT 2007), pp.97-104 (2007).
[12] D. Noguchi, S. Sugimoto, Y. Bannai, and K. Okada, "Time Characteristics of Olfaction in a Single Breath", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.83-92 (2011).
[13] A. Aruga, Y. Bannai, and T. Seno, "Investigation of the Influence of Scent on Self-Motion Feeling by Vection", International Journal of Informatics Society, Vol.11, No.2, pp.65-73 (2019).
[14] Q. Yan, J. You, W. Sun, Y. Wang, H. Wang, and L. Zhang, "Advances in Piezoelectric Jet and Atomization Devices", Applied Sciences, Vol.11, No.11, p.5093 (2021).
[15] S. Nakamura, Y. Bannai, "Development of An Olfactory Display Using a Piezoelectric Element and Measurement of Olfactory Detection Threshold", VRSJ Research Report, Vol.25, No.SBR-1, pp.1-6 (2020). (in Japanese)
[16] Y. Seta, M. Makino, Y. Bannai, and M. Hattori, "Proposal of Air-Flow Channel Structures of The Olfactory Display Using the Piezoelectric Element Considering Fluid Behavior", VRSJ Research Report, Vol.27, No.SBR-1, pp.9-14 (2022). (in Japanese)
[17] M. Tachibana, and K. Yoshida, "Energy Loss due to Honeycomb in a Pipe Flow", Memoirs of the Faculty of Engineering, Fukui University, Vol.43, No.1, pp.103-108 (1995). (in Japanese)
[18] Y. Kohama, "Optimum Screens for Turbulence Reduction in Wind Tunnels", Journal of Japan Society of Fluid Mechanics Nagare, Vol.9, No.1, pp.19-33 (1990). (in Japanese)
[19] Y. Seta, N. Mori, M. Makino, Y. Bannai, and M. Hattori, "Comparison of Rectification Effects of Grid Shapes in an Air-Flow Channel for Droplet Spray Olfactory Displays", Proceedings of the 84th National Convention of IPSJ, p.2F-05 (2022). (in Japanese)

(Received: February 1, 2023)
(Accepted: July 5, 2023)

Yohei Seta He is a Doctor Course Student of Graduate School of Science and Engineering at Chuo University. He received a Bachelor of Media Science and a Master of Media Science from Tokyo University of Technology. in 2009 and 2011, respectively. His research interests include real-time computer graphics, virtual reality, computational fluid dynamics.

Mitsunori Makino received his B.E, M.E and Dr. Eng in Electronic Communication Engineering from Waseda University in 1987,1989 and 1992, respectively. He was a research assistant at Waseda University from 1991 to 1992. He joined Chuo University in 1992, and has been a full professor of the Faculty of Science and Engineering, Chuo University since 2004. His research interests are mainly in technologies and their applications in CG, XR and visualization. His activities also include engineering education. He is currently the director and the committee chairperson of the Criteria in Japan Accreditation Board of Engineering Education.

Yuichi Bannai He is a professor of Department of Information Media at Kanagawa Institute of Technology, Japan. He received a BE and a ME from Waseda Univ. in 1978 and 1980, respectively, and joined Canon Inc. in 1980. He also received MS from Michigan State University in 1988, and Ph D from Keio University in 2007. His research interests include human five senses interaction, artificial intelligence, and virtual/augmented reality. He is a member of ACM, IEEE CS, IPSJ, Japanese Society for AI, and Virtual Reality Society of Japan (VRSJ).

Motofumi Hattori He is a professor of Department of Information Media at Kanagawa Institute of Technology, Japan. He received a Master of Science from Kobe University in 1991. He received a Master of Engineering from Kobe University in 1993. He received a Doctor of Engineering from Kobe University in 2000. His research interests include Game Amusement, Virtual Reality, Computational Fluid Dynamics, and Statistical Physics. He is a member of Game Amusement Society of Japan, Virtual Reality Society of Japan, IPSJ, The Japan Society for Computational Engineering and Science, and The Japan Society of Fluid Mechanics.

Regular Paper

Effectiveness and Issues of Self-assessment for Empowerment Arts Therapy Practitioners

Michi Komura†, Midori Ishihara‡, Hajime Kaneko*, and Akihiro Hayashi**

†Osaka Metropolitan University, Japan
‡ATAS Laboratory, Japan
*Kobe University of Future Health Sciences, Japan
**Shizuoka Institute of Science and Technology, Japan
{mitikomm347, midorippa2021, kaneko68h, pixysbrain}@gmail.com

Abstract -To improve the reliability of Empowerment Arts Therapy (EAT) conducted in local communities and to increase the benefits to its users, there is a need for continuous improvement in the capabilities of EAT practitioners. The authors have proposed a self-assessment method that consists of several assessment tools for this purpose. It is suggested that these self-assessment tools are effective to deepen practitioners' self-recognition of their EAT activities and promote their motivation toward self-improvement. In this paper, we discuss and report the effectiveness and issues of self-assessment for EAT practitioners based on the results of a trial case using self-assessment tools.

Keywords: Empowerment Arts Therapy, Self-assessment, Self-improvement, Community Formation, Mental Health.

1 INTRODUCTION

Empowerment Arts Therapy in the community is an activity that uses artistic techniques such as painting, music, drama, dance, and poetry to support people's lives, develop their capabilities, and improve their quality of life (QOL) in empowering ways. The main target of EAT includes people in relatively good health. EAT is distinguished from Psychopathological Arts Therapy (PAT) which is psychotherapy for the treatment of mental illness. EAT has a wider variety of forms and methods than PAT and is not subject to established methodologies or public institutional frameworks.

EAT practitioners have a wide range of backgrounds and skill levels. The practitioners do not necessarily aim at therapeutic activities from the beginning of their careers. They are engaged in various artistic activities such as painting classes, music classes, dance groups, and so on. Many practitioners independently learned arts therapy during the process of seeking solutions to improve the mental problems of their clients and developed their unique style of activities. This diversity makes it difficult to uniformly evaluate the skills of EAT practitioners. Evaluation criteria have not yet been established. It is, however, necessary to clarify and disseminate the basic requirements guiding EAT practices. This is to ensure that the practices are safely conducted and avoid causing unintentional disadvantages because EAT activities are community-based and involve people's mental health issues. Furthermore, it is essential to continuously improve EAT practitioners' capabilities based on the basic requirements.

In this paper, we outline self-assessment tools including the EAT Assessment Sheet and the EAT Rubric for practitioners to improve their capabilities as a solution to the lack of evaluation criteria. We examine a trial case for the self-assessment tools, which have been revised and improved based on the results of previous surveys. Additionally, we clarify the effectiveness of the self-assessment for EAT practitioners and highlight future issues. Section 2 presents an outline of the self-assessment tools and the purpose of the study. Section 3 describes the expected effects of the self-assessment. Section 4 reports on the results of a trial case in an EAT practitioner. This trial case study is showcased as preliminary research before the full-scale verification study to be conducted in the future. Section 5 describes the effectiveness of self-assessment and future issues.

2 OUTLINE OF THE SELF-ASSESSMENT TOOLS AND PURPOSE OF THE STUDY

2.1 Outline of the Self-Assessment Tools

EAT is the concept defined and proposed by Arts Therapy Activities Study (ATAS)[1] which we have been engaged in since 2012. PAT is primarily conducted in clinical settings by psychotherapists and psychiatrists. In recent years, activities utilizing arts therapy have been expanding beyond the clinical settings in Japan. We can find cases of practice such as a trial of the technique "picture drawing play" with an application of psychotherapy in homeroom activity in elementary school [1], workshops conducted by a cloth collage artist [2], and mosaic-making session process of a contract office worker in adult mosaic classes conducted by a mosaic artist [3]. In the trial case of practice in an elementary school, direct practitioners were schoolteachers, not psychotherapists, although the project was conducted in collaboration between

[1] In ATAS, two research projects "National Survey on Arts Therapy" (2012-2014) and "Elucidation of the Constitutive Requirements and Development of Evaluation Criteria for Empowerment Arts Therapy" (2015-2017) were conducted being funded by the Grant-in-Aid for Scientific Research by Japan Society for the Promotion of Science (JSPS).

a psychologist and schoolteachers. In the cases of cloth collage and mosaic production, the practitioners were both artists. Furthermore, the interdisciplinary study conducted at the Konan Institute of Human Sciences (KIHS) revealed the expansion of arts therapy activities in everyday life [4]. These activities utilize arts therapy but are conducted for different purposes than clinical treatment. These activities can be positioned in the area of EAT. In line with this trend, the practitioners involved in these activities is expanding from medical or clinical professionals to people with a broader range of background.

Thus, EAT activities are socially expanding in Japan. However, the capabilities of EAT practitioners have not been discussed. Specific methods for their assessment and improvement have not been established.

In the area of PAT, there are some discussions about arts therapy assessments. Lefèvre et al. [5] discussed the impacts of arts therapy on palliative cancer patients' distress reduction based on self-assessments by the patients. Beard's systematic literature review [6] includes analyses regarding assessment methods and tools for various kinds of arts therapy for dementia patients containing self-assessment by the patients. Betts [7] broadly examined art therapy assessments and rating instruments and discussed their applicability, validity, and reliability. These studies target arts therapy for patients in clinical settings with the primary purpose of treating diseases or relieving symptoms and describe benefits and issues for arts therapy assessments including self-assessment by patients themselves. On the other hand, EAT is conducted for the general public in communities outside the clinical settings. The circumstances and needs of EAT vary from those of PAT. Assessment methods suitable to EAT circumstances and challenges will be required. Unlike in PAT, however, there have been few scales and little discussion about assessment methods fit for EAT practices and practitioners.

The development of a self-assessment tool for EAT practitioners was derived from ATAS. In 2012, ATAS conducted a nationwide survey project[2] on various kinds of arts therapy activities in non-clinical fields [8], and elucidated the basic constructive requirements of EAT [9]. Through this nationwide survey, the necessity of capability evaluation of EAT practitioners was identified. Self-assessment was proposed as a solution to this issue. It was because establishing an objective assessment method at a leap was considered not feasible given the actual environment around EAT [9] (pp.9-10).

However, the implementable assessment tools had not yet been created. Therefore, the EAT Assessment Sheet (Appendix I) was created to provide a practical assessment tool in a user-friendly form [10].

The EAT Assessment Sheet intended to help EAT practitioners understand their own strengths and weaknesses, and to make use of their efforts toward self-development. However, there was a risk that the criteria may be too subjective when the sheet alone was used. In addition, it was difficult for EAT practitioners to recognize what kind of state to aim for in the next stage to improve items with low scores. To solve this issue, the EAT Rubric was created as a complementary tool to the EAT Assessment Sheet [11].

A rubric is "a scoring tool that lays out the specific expectations for an assignment" [12]. It contains scales, dimensions, and descriptions of the dimensions for a task. In recent years, rubrics have been applied from elementary to higher education because interactive and formative assessment has become more important as Japanese education has shifted to learner-centered active learning.

Rubrics are used not only for assessment by teachers but also for self-assessment by students. Hoshi and Koshikawa [13] analyzed the twenty-six papers concerning the use of rubrics for students' self-assessment in university education and suggested the effects of self-assessment with rubrics: students' understanding objectives, motivation towards learning, improvement of qualities and abilities, and teachers' understanding students. Rubrics are considered to contain elements that promote learners' progress as well as assessment functions. These features of rubrics are expected to be effective enough for self-assessment in the field of social activities such as EAT.

In addition to the sheet and rubric, a mapping sheet was prepared for EAT practitioners to check and recognize the positions of their activities. The mapping sheet is adapted from a conceptual diagram originally created to explain the interrelationships and positions of concepts and activities such as arts therapy, art, and empowerment [14] (p.106).

The basic components of the self-assessment tools at the present stage are described below, though the development and revision of each tool are ongoing.

(1) Mapping Sheet (Appendix I)

This is a map for EAT practitioners to grasp the position of their activities by referring to the Five Types of Arts Therapy and the difference between EAT and PAT. EAT practitioners mark the areas where their activities overlap on the diagram of "Positioning of Art/Therapy/Empowerment" in the mapping sheet. This work promotes reflection on their activities and visualizes their positioning. Mapping is expected to serve as an introduction or preparation for self-assessment.

(2) EAT Assessment Sheet (Appendix II)

This is a checklist-type sheet consisting of 26 items in 8 categories related to the skills and knowledge of EAT practitioners, their psychological safety, sustainability, and the basic principles underlying EAT activities. The content of each item is based on the constructive requirements for EAT as elucidated by ATAS [9] (pp.11-18). EAT practitioners rate each item on a 4-point scale, calculate the average values[3] for

[2] "National Survey on Arts Therapy" (2012-2014) funded by the Grant-in-Aid for Scientific Research by Japan Society for the Promotion of Science (JSPS).

[3] The numerical value of each item is an ordinal scale, and the calculation of the average value does not have statistical effectiveness. However, it is important for the items to be easy to understand and simple to use by EAT practitioners to grasp their own status and position. Based on this understanding, this study followed the examples of quantification in social psychological tests such as vocational aptitude and personality tests. In these tests, the ordinal scale is regarded as an interval scale and was quantified, which is useful in grasping one's position in an overview using a ranking process. However, the appropriateness of using averages in the EAT self-

each category, and subsequently draw a radar chart. This work helps visualize their skill levels and enables EAT practitioners to recognize the actual conditions and characteristics of their activities.

(3) EAT Rubric (Appendix III)

The EAT Rubric is based on the rubric format as a criterion for self-assessment regarding EAT. It has seven perspectives: Knowledge, Skills, Psychological Safety, Self-exploration /Self-understanding, Cooperation/Collaboration, Management/Sustainability, and Basic Principles. The EAT Rubric has the following four-level scale for each perspective:

1. A state that should have been achieved. The items checked here indicate that there is much room and need for improvement.
2. A state that indicates inadequacy, but the practitioner is on the way to improvement.
3. A state that generally meets expected standards.
4. A state that serves as a goal or guideline for each item.

It is, however, difficult to specify objective criteria such as which level or higher is acceptable and which level or lower is not because the evaluation criteria for the EAT are still under research. The EAT Rubric is not intended to be used as a stand-alone tool, but rather as a reference of criteria for self-assessment with the EAT Assessment Sheet. For these reasons, the rubric does not contain evaluative terms such as excellent, good, and inadequate that would make EAT practitioners aware of passing or failing, instead, numbers from 1 to 4 are used.

The EAT Rubric was created in consideration of the correspondence with the items of the EAT Assessment Sheet. The eight categories on the EAT Assessment Sheet were reorganized into seven assessment dimensions to cover the important points concisely although the rubric dimensions on the EAT Rubric do not completely correspond to the categories on the EAT Assessment Sheet. The description of the EAT Rubric is more abstract than that of the EAT Assessment Sheet which lists specific items. The EAT Rubric provides a comprehensive bird's-eye view of the requirements for EAT while the EAT Assessment Sheet provides concrete checking items. By using the assessment sheet and the rubric in a complementary manner, it is expected to promote a bird's-eye view and specified understanding of the status of one's own activities and skills. This is why the two kinds of tools are separately prepared to be used together.

On the other hand, there might be a risk that using the two different tools together may complicate the process and the usage and purpose of each tool may not be well understood. This risk can be avoidable by providing concise and easy-to-understand instructions and guides, which is one of the upcoming tasks.

assessment remains to be analyzed and needs to be carefully examined in the course of future studies [15].

2.2 Purpose of the Study

The fundamental issue in EAT is how to ensure the capabilities of practitioners and the quality of activities from among a diverse range for which there are no established evaluation criteria. The purpose of this study is to examine the effectiveness of continuous self-assessment using the self-assessment tools described above.

A previous survey conducted before this study suggested a certain level of effectiveness for self-assessment using the EAT Assessment Sheet in combination with the EAT Rubric [16]. It is expected that EAT practitioners will deepen their self-awareness and increase their motivation for autonomous self-improvement by self-reflection in a fixed-point manner through regular self-assessment work.

This study examines the effectiveness and issues concerning self-assessment based on a case study of the trial use of self-assessment tools by an EAT practitioner.

3 EXPECTED EFFECTIVENESS OF SELF-ASSESSMENT

3.1 Continuous Self-improvement

The expected effect of self-assessment is the self-improvement of EAT practitioners through its continuous and regular implementation. The objective of the self-assessment tools described in Chapter 2 includes supporting and encouraging EAT practitioners in their efforts to increase the quality and reliability of their activities by objectively viewing their own activities and skill levels. Further, the tools aim at helping to independently determine what characteristics to develop as strengths and those that require improvement by regular monitoring.

It should be emphasized that the main actors undergoing assessment and improvement are EAT practitioners themselves. The self-assessment tools are mechanisms to help empower them in their efforts and are not intended to draw a line between pass/fail or suitable/unsuitable regarding the quality of activities and skills of EAT practitioners based on uniform criteria.

3.2 Improvement Process

The following process for improvement of practitioners' capabilities is envisioned by applying the self-assessment tools (Fig. 1):

1) Clarify areas of activities using the mapping sheet (To be necessarily conducted the first time. Thereafter, it should be performed as required).
2) Self-assess with the EAT Assessment Sheet alone (To be conducted the first time only*).
3) Self-assess with the EAT Assessment Sheet referring to the EAT Rubric**.
4) Identify the area(s) that require improvement and list the specific item(s) to be improved.

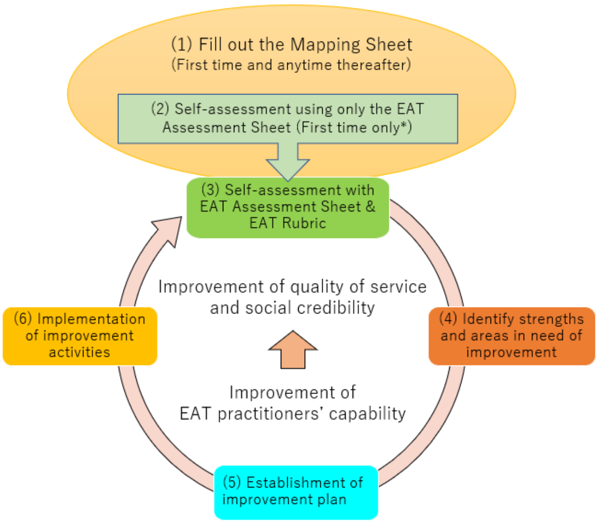


Figure 1: Improvement Process

5) Formulate an improvement plan. In this process, it is essential to carefully examine the improvement items listed in light of one's own activity policy, prioritize these items, and formulate a concrete and feasible plan.
6) Implement improvement activities according to the plan.
7) Reassess as described in point 3 above.

*In the first self-assessment session, processes 2) and 3) should be conducted, and the differences between both results should be reviewed.

**From the second time onward, the process from 3) to 7) should be conducted to form a baseline assessment. Process 1 should be conducted as required, such as when there is a change in areas of activities.

4 TRIAL CASE FOR SELF-ASSESSMENT

This chapter presents a case study in which an informant of an EAT practitioner conducted self-assessments at specific intervals.

4.1 Outline of the Survey

4.1.1 Informant: Practitioner A

- Instructor of art classes
- Number of students: Approx. 120
- Self-employed/Female/Age in the 50s
- Approx. 28 years of experience as an art instructor
- Approx. 8 years of experience in EAT activities.

Practitioner A conducts various kinds of artistic activities mainly aimed at youth in her art studio in the local community. Initially, she worked on a volunteer basis, but as the scale of her activities expanded, she turned the activities into a private business. She came to recognize the necessity and effectiveness of EAT through her painting classes and has been studying EAT on her own initiative to improve herself as an EAT practitioner. She has established a network with local welfare institutions and psychiatric medicine departments.

The characteristics of Practitioner A's activities are as follows.

- Her art classes are conducted as general painting classes. They do not target children with special needs alone.
- Individual EAT activities are conducted with children who have specific needs, if any, in the painting class activities.
- The practitioner relates to children in an EAT-like manner during the usual art class activities.

4.1.2 Method of Investigation

1) Trial I: The informant completed the self-assessment in two ways:
   i. Trial I-i: Self-assessment was conducted using the EAT activity evaluation sheet alone without the EAT Rubric (Method A).
   ii. Trial I-ii: Self-assessment was conducted using the EAT Activity Assessment Sheet referring to the EAT Rubric (Method B).
2) Trial II: Approximately three months after Trial I, a second self-assessment trial was conducted. The Mapping and Review Sheets were newly added based on the discussion at the time of Trial I. The method of Trial II was as follows.
   i. Mark areas of activities on the Mapping Sheet.
   ii. Self-assess using the EAT Assessment Sheet and EAT Rubric.
   iii. Fill in a review sheet[4].
3) The informant was interviewed about her perceptions and changes after performing Trials I and II.

4.2 Results of the Trial Case

4.2.1 Visualization of Position by Mapping

Practitioner A marked the different activity areas according to the activity targets (e.g., children, adults) and the form of each activity (individual, group) (Fig. 2). Her Mapping Sheet indicates that she covers the areas that span from EAT to art. The results suggest that Practitioner A has been expanding the scope of her activities over the course of her career.

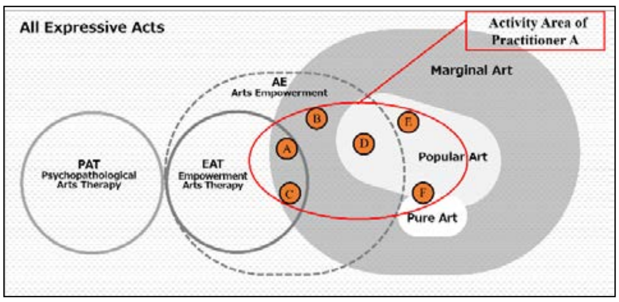


Figure 2: Practitioner A's Mapping Sheet

[4] It is a sheet for EAT practitioners to write their findings and reflections through self-assessment.

(A) Individual session for children with specific needs
(B) Children’s class in a group
(C) Illustration circle for youth
(D) Adult class
(E) Course for citizens (e.g., in a community center)
(F) Artwork creation of her own

The Mapping Sheet was added based on discussion and reflection after Trial I. It was ascertained that it would be desirable to have a process to clarify one’s activity areas before performing the self-assessment using the EAT Assessment Sheet and EAT Rubric in Trial I (4.1.2). The process of mapping provides an opportunity to look at one’s activities from a bird’s-eye view and is effective for preparation and introduction to the subsequent self-assessment.

4.2.2 Changes through Self-assessment

Practitioner A’s radar charts based on the results of a series of self-assessment trials are shown in Fig. 3 to Fig. 5.

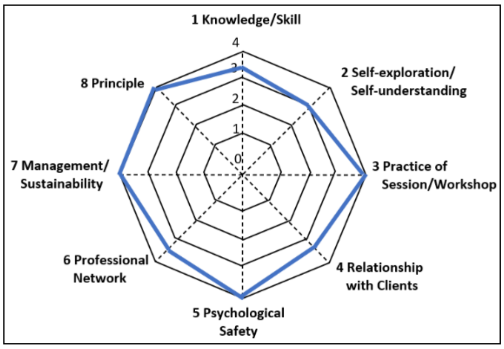


Figure 3: Trial I-i (Aug. 2022)

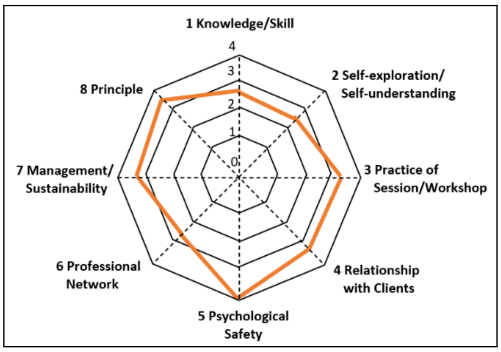


Figure 4: Trial I-ii (Aug. 2022)

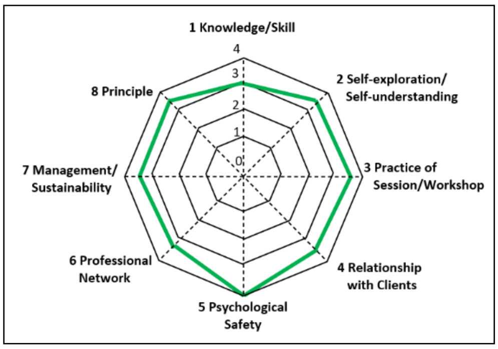


Figure 5: Trial II (Nov. 2022)

The difference between Trial I-i (Fig. 3) and Trial I-ii (Fig. 4) can be attributed to the presence or absence of the EAT Rubric. Trials I-i and I-ii were conducted at the same time. Trial I-ii was assessed more severely than Trial I-i which was performed without reference to the EAT Rubric. It can be inferred that the presence of the EAT Rubric let the practitioner aware of external criteria and cause some change in her perspective of self-assessment. It might be controversial whether severe self-assessment directly leads to the improvement of capability. Practitioner A, however, stated in the interview after Trial I-ii that the EAT Rubric helped her better understand the criteria and made self-assessment easier.

Comparing the results of Trial I-ii (Fig. 4) and Trial II (Fig. 5), the overall balance of figures has improved in Trial II. Based on the results of the self-assessment after nearly three months, an interview with Practitioner A was conducted to examine what she noticed, and how her awareness, vision, and actions toward improvement changed after the previous trials.

The feedback from Practitioner A was as follows.

Feedback from Practitioner A

- Through working with the EAT Assessment Sheet and the EAT Rubric, I became aware of things that I had not been aware of regarding my activities.
- The Mapping Sheet helped me to organize and get a bird’s-eye view of my activities that have expanded over the years.
- Self-assessment gave me an opportunity to reflect on myself. Until then, I had been rushing through the process using trial and error to meet the needs which I was asked in the community but had never looked back over my activities objectively.
- I had been thinking of retiring in 10 years because of my age. However, through the self-assessment process, I have been able to reflect on my past and reperceived the needs and the social necessity of my activities. Considering that I am needed in the community, my thoughts about retirement are wavering. I now feel like continuing to work as long as I have the vigor to do so.
- I have tried to train and develop successors, but I find it difficult due to financial problems and the communication skills required for practitioners. I am thinking of limiting ourselves to private classes alone 10 years from now. Staff and finances are continuous issues related to the sustainability of activities.
- I have learned a variety of methods and pursued a unique method for my activities. On the other hand, I have a sense of regret about the fact that I have been doing this without much interaction with others. Now I strongly feel that I need to relearn and brush up.
- The self-assessment tools give me a third-party perspective. I think it is meaningful to look at our activities not only from our own perspective but also from a social viewpoint to think of the future of our activities.

(Based on an interview and review sheets)

5 EFFECTIVENESS AND FUTURE ISSUES

5.1 Effectiveness of Self-Assessment

Based on the discussion of Practitioner A's trial case, the following effects of self-assessment were identified:

1) The process of mapping is useful to look at one's activity area from a bird's-eye view and to grasp the characteristics of each activity.
2) Through self-assessment using the EAT Assessment Sheet and the EAT Rubric, EAT practitioners become aware of external criteria and increase awareness. This will naturally change the perception of the status of their activities and future outlook. It can lead to motivation for learning.
3) The two-step process in the initial self-assessment, i.e., checking the EAT Assessment Sheet without referring to the EAT Rubric for the first time and then referring to it helps EAT practitioners to be aware of external criteria and leads to shift their perspectives of self-assessment. One of the expected effects of the rubric is the improvement of metacognitive ability [17] [18]. EAT Rubric is considered to include the potentiality to develop EAT practitioners' ability for self-assessment.
4) Mapping work and radar charts are useful for visualizing the current state of EAT practitioners at a given time and for observing changes over time.

From the above, we concluded the self-assessment tools could help EAT practitioners recognize their activities and skills by themselves and had the potentiality to promote the deepening of their self-recognition.

5.2 Future Issues

Firstly, this study is based on trials by one informant over a period of three months, which is insufficient to generalize the above effects. The informant in this case study has several years of experience and is thought to have already reached a considerably high level of proficiency regarding the items indicated in the EAT activity Assessment Sheet and the EAT Rubric through trial and error. Therefore, it was not possible to determine to what extent a clear change or improvement in her skill level was observed before and after the self-assessment with this series of tools. Larger-scale and longer-term studies involving a wider variety of practitioners as informants are needed in order to verify the effectiveness of these self-assessment tools.

Secondly, a problem with the EAT self-assessment process is the absence of perspectives other than EAT practitioners. In the educational field, the presence and intervention of teachers is a basic premise. Even self-assessment is supervised by teachers and proper feedback is provided to students. Uchiyama and Itoh [18] described the importance of others' views to utilize the assessment with rubrics in designing courses. Nishikata [19] pointed out the importance of teachers' feedback in promoting students' autonomous motivation. In the field of education, the interaction between students as self-assessors and teachers as supervisors or facilitators leads to the development of students' cognitive abilities and skills and the improvement of courses and classes. However, the EAT self-assessment process does not contain such supervisors or facilitators. It is because EAT practitioners independently conduct their EAT activities in many cases, which are commonly grass-root and voluntary. This is an important issue to ensure the validity of the criteria and results of the self-assessment.

This study has the significance of having concretely provided a concrete method for EAT practitioners who had almost no indicators to refer to despite these limitations. This study will be a base or a starting point for further future discussion to improve the capabilities of EAT practitioners and the quality of activities.

Finally, we would like to mention the relationship between this research and Information and Communication Technology (ICT).

EAT activities have so far owed much to the dedicated efforts of EAT practitioners. Since most of the practitioners independently conduct EAT activities on a volunteer basis, they have few appropriate tools to refer to the best practices required for their activities.

They make efforts to grasp their clients' needs, recognize the progress and measure the effectiveness and influence of their EAT activities in their own way. Nevertheless, it is not easy for EAT practitioners to record and analyze the process of their activities to reflect the findings in future performance improvement, partly due to the lack of financial resources to cover such efforts. In addition, EAT practitioners are sometimes not good at ICT. Therefore, EAT activities have been alienated from ICT so far.

In this study, EAT Assessment Sheet and EAT Rubric were created in a digitally distributed format. These two tools enable EAT practitioners to quickly record and accumulate data. When EAT activity data are accumulated enough in a standardized format, they allow the rational evaluation of comparative discussion among plural EAT practitioners.

If EAT practitioners rely solely on their own experience to carry out EAT activities, it may lead to unfavorable results such as neglecting necessary basic movements. These two tools can be beneficial for EAT practitioners to make a habit of PDCA in daily activities and to reduce their wasted effort in PDCA. This study can be positioned as a preparatory process for the utilization of ICT in EAT activities.

One of the future issues is to create a web-based system that can accumulate and analyze EAT practitioners but requires as little cost of development and maintenance as possible.

REFERENCES

[1] T. Okada, "Psycho-educational Approach with Application of Psychotherapy in Elementary School - Trial and Results of the Technique: 'Picture Drawing Play'", Japanese Bulletin of Arts Therapy, Vol.40, No.1, pp.43-51 (2009) (in Japanese).
[2] T. Fujii, "A Postcard Artist's Report on Cloth Collage Workshops", Japanese Bulletin of Arts Therapy, Vol.41, No.1, pp.70-80 (2010) (in Japanese).
[3] N. Okada, "Mosaic, Reflections in Art Therapy", Japanese Bulletin of Arts Therapy, Vol.50, No.1, pp.90-97 (2019), (in Japanese).

[4] M. Ishihara, “Nichijo ni yorisou art therapy - Kosodate-shien toshite no katsudo-jirei ni art therapy no hattenkei wo miru -”, K. Tokiko, and K. Nishi (eds.), Art Therapy Saiko - Geijutsugaku to rinsho no genba kara -, Heibonsha, Tokyo, pp.232-253 (2013) (in Japanese).

[5] C. Lefèvre, M. Ledoux, and M. Filbet, “Art Therapy among Palliative Cancer Patients - Aesthetic Dimensions and Impacts on Symptoms -”, Palliative and Supportive Care, Vol.14, No. 4, pp.376-380 (2016).

[6] R. L. Beard, “Art Therapies and Dementia Care - A Systematic Review -”, Dementia, Vol. 11, No. 5, pp.633-656 (2012).

[7] D. J. Betts, “Art Therapy Assessments and Rating Instruments - Do They Measure Up? -”, The Arts in Psychotherapy, Vol. 33, No. 5, pp.422-434 (2006).

[8] M. Ishihara, and H. Kaneko, “Empowerment toshite no shisei no art therapy katsudo - Zekoku jittai-chosa kara mieru sono naihatsusei to jiritusei -”, The Crisis of the Mind and Clinical Knowledge, Vol.16, pp.105-130 (2015) (in Japanese).

[9] H. Kaneko, and M. Ishihara, “A View of ‘Art Therapy for Empowerment’”, Journal of the Kobe University of Welfare, Vol.19, No.1, pp.1-22 (2018) (in Japanese).

[10] M. Komura, and A. Hayashi, “A Method of Assessment for Empowerment Arts Therapy”, IEICE Technical Report, Vol.121, No.372, pp. 1-7 (2022) (in Japanese).

[11] M. Komura, M. Ishihara, and A. Hayashi, “Proposal for Continuous Improvement of Empowerment Arts Therapy - A Method of Self-Assessment by Applying Rubric -”, Proc. International Workshop of Informatics (IWIN 2022), pp.121-126, Nachi-Katsuura, Japan (2022).

[12] D. Stevens, and A. Levi, Introduction to Rubrics, Second Edition, Stylus Publishing, VA, Chapter 1 (2013).

[13] Y. Hoshi, and S. Koshikawa, “Effects and Challenges of Using Rubrics for Self-Assessment in University Education”, Japan Bulletin of Educators for Human Development, Vol.23, No.1, pp.21-30 (2020), (In Japanese).

[14] M. Ishihara, “Nichijo ni nezasu art to art therapy - “Empowerment” gainen ni yotte miete kuru kozu -”, a+a Bigaku Kenkyu (Studies in Aesthetics and Art Criticism), Vol.10, pp.96-109, p.106 (2017) (in Japanese).

[15] M. Ishihara, M. Komura, A. Hayashi, and H. Kaneko, “Empowerment gata arts therapy no self-assessment no jitsuyoka ni mukete - Empower suru hito wo empower suru tameni -”, The Crisis of the Mind and Clinical Knowledge (Bulletin of Konan Institute of Human Sciences, Konan University), Vol. 24, pp.1-21 (2023) (In Japanese).

[16] M. Komura, M. Ishihara, H. Kaneko, and A. Hayashi, “Effectiveness Applying a Rubric in Self-Assessment of Empowerment Arts Therapy Practitioners”, IEICE Technical Report, Vol. 122, No. 282, pp. 33-39 (2022) (in Japanese).

[17] K. Terashima, and T. Hayashi, “Development of Problem-Based Learning to Promote Self-Evaluation Using Rubrics”, Kyoto University Researches in Higher Education, Vol.12, pp.63-71 (2006) (in Japanese).

[18] C. Uchiyama, and M. Itoh, “How Do Science Teachers Design Lessons Using of Rubrics Effectively for Assessment? - An Investigation of Junior High and High School Science Education -”, Journal of Science Education in Japan, Vol.42 No.1, pp37-47 (2018) (in Japanese).

[19] H. Nishikata, “The Effects of Designing a Rubric and Self-Evaluation - Focusing on Internalization of Autonomous Motivation -”, Japan Journal of Educational Technology, Vol.43, No.3, pp.215-229 (2019) (in Japanese).

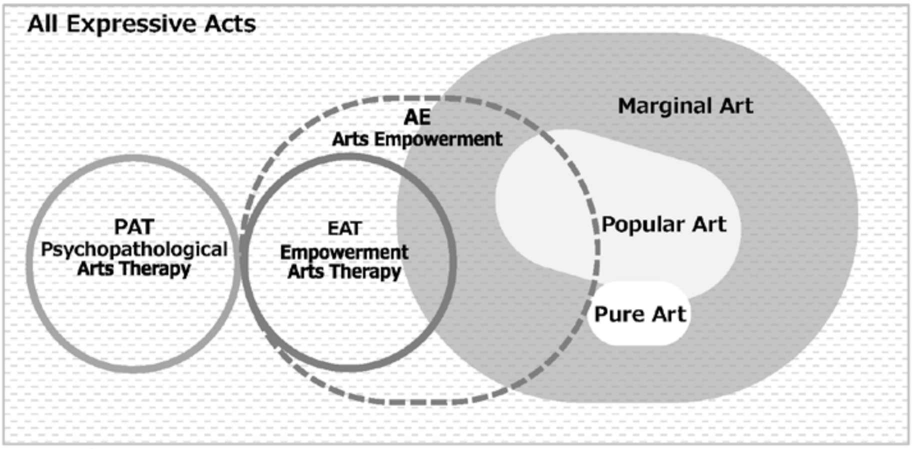
(Received: March 3, 2023)
(Accepted: June 19, 2023)

APPENDICES

Appendix I Mapping Sheet

Referring to the *Five Types of Arts Therapy* and the *Difference between PAT and EAT*, mark the areas covered by your activity on the diagram below for ***Positioning of Art/Therapy/Empowerment***.

[Positioning of Art/Therapy/Empowerment]



***AE: Arts Empowerment**
Art activities that revitalize communities and towns such as community art, and art activities that unintentionally have various unintentional therapeutic effects. Both professional and amateur artists are included.

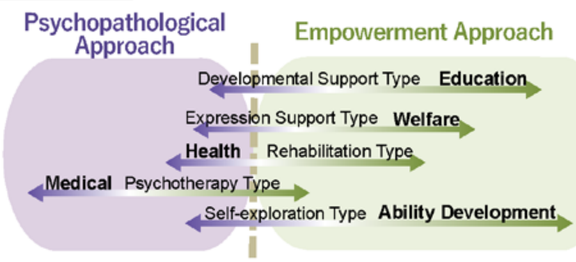
***Marginal Art/Popular Art/Pure Art**
The concepts were proposed by TSURUMI, Shunsuke (1922–2015), a Japanese philosopher. **Marginal Art** is created and enjoyed by non-professionals, forming a vast area where art and life permeate each other. It is said to be "the first form of art" and to "have the power to give birth to pure art and popular art". **Pure art** is created by professionals and enjoyed by connoisseurs who are versed in these fields. **Popular art** is created by professionals or corporations and enjoyed (consumed) by non-specialists = the masses.
TSURUMI Shunsuke, "*Genkai-geijutsu ron*" [*Theory of Marginal Art*] (1991, Chikuma Gakugei Bunko [first published in 1967, Keiso Shobo])

Five Types of Arts Therapy

Psychotherapy	Targets persons in need of diagnosis and treatment by experts. Utilized as a major or adjunctive means of alleviating or treating problems and symptoms. There is a lot of overlap with self-exploration Type in what is aimed to achieve.
Self-exploration	Targets adults with relatively good health. The main purposes are distraction, stress reduction, recreation, healing, self-liberation, self-discovery, self-affirmation, and quality of life improvement. Other purposes include confronting problems and preventing mental illness. It may lead to Psychotherapy Type.
Rehabilitation	Targets people with brain dysfunction and the elderly. The main purpose is to restore physical function and control, alleviate, and stabilize the progression of symptoms. Utilized as an aid or traction for rehabilitation. It differs from psychotherapy, although it includes psychological aspects.
Developmental Support	The program targets children and adolescents in their developmental stages. The purposes are to free children from oppression, promote self-expression, increase self-confidence, and improve self-esteem. For children with various disabilities, it is a means of "rehabilitation". All of these include elements of parenting support. While it may lead to psychotherapy, it differs from psychotherapy.
Expression Support	Targets persons who have difficulty in normal communication and social life due to intellectual or mental disabilities, or people who engage in spontaneous and intrinsic expressive activities as a revelation of these symptoms. The main purpose is to improve the quality of life by enabling them to interact and spend quality time with the outside world through non-verbal expressive acts. It differs from psychotherapy, although it involves psychological aspects.

Difference between PAT and EAT

The condition of a person concerned is viewed as abnormal and in need of treatment. The main principle is to treat and normalize their condition according to the treatment purpose. The patient/client can choose the method of treatment, but if he/she hopes to not receive treatment, that choice tends to be difficult or unacceptable.



Long term support is provided so that the person concerned can discover, develop, and show his/her potential, abilities, and motivation. He/she will not be forced to conform him-/herself to conventional wisdom commonly accepted in the society.

Therapist → Patient Client

Supporter ⟷ Person Concerned

Appendix II EAT Assessment Sheet

I CAPABILITY OF PRACTITIONERS

1 Knowledge and Skills

		Scoring	Average
(1)	Understand and able to explain the therapeutic effects and functions of art and expression.	1・2・3・4	Sum of (1) to (4)
(2)	Possess knowledge and skills in arts therapy and art to conduct sessions and workshops that meet their purposes and aims.	1・2・3・4	
(3)	Know the history of arts therapy, techniques, and characteristics of arts therapy outside of your own area of expertise.	1・2・3・4	
(4)	Possess knowledge and skills in psychotherapy and psychological counseling at the level required in your own activities.	1・2・3・4	Average

2 Self-improvement/Self-understanding

(5)	Work on improving knowledge and skills required for your expertise and related areas for activities.	1・2・3・4	Sum of (5) to (7)
(6)	Conduct (periodic) self-assessment.	1・2・3・4	
(7)	Understand your nature and propensities, and your own psychological issues have been resolved to a certain degree.	1・2・3・4	Average

II PRACTICE

3 Practice of Sessions and Workshops

(8)	The purposes, aims, and targets (clients) of your activities are clear.	1・2・3・4	Sum of (8) to (11)
(9)	Able to set goals, construct, and conduct sessions and workshops that meet the clients' condition and needs.	1・2・3・4	
(10)	Able to conduct sessions and workshops in a flexible manner according to the clients' condition (Change or suspend sessions or workshops if necessary).	1・2・3・4	
(11)	Measure the effectiveness of the sessions and provide feedback.	1・2・3・4	Average

4 Relationship with Clients

(12)	Able to establish an appropriate relationship with a client, taking into account the required psychological distance between the subject and the company.	1・2・3・4	Sum of (12) to (14)
(13)	Strive to understand clients from multiple perspectives.	1・2・3・4	
(14)	Strive to obtain information about the clients' family environment, social environment, and relevant institutions, organizations and services.	1・2・3・4	Average

5 Psychological Safety

(15)	Maintain confidentiality of clients' personal information and their words and deeds.	1・2・3・4	Sum of (15) to (17)
(16)	Confirm that participants should maintain confidentiality of each participant's personal information and words and deeds when conducting group work or sessions.	1・2・3・4	
(17)	Provide necessary individual follow-up depending clients' psychological condition and changes.	1・2・3・4	Average

6 Professional Network

(18)	Have a supervisor or a mentor to consult with.	1・2・3・4	Sum of (18) to (20)
(19)	Able to connects clients to an appropriate expert or expertise agencies depending on the clients' condition and the nature of the case.	1・2・3・4	
(20)	Possess a network to obtain information and knowledge related to arts therapy and your own activities.	1・2・3・4	Average

III MANAGEMENT STYLE

7 Management/Sustainability

(21)	Make income and expenditure plans and carry out continous activities, or aim to do so.	1・2・3・4	Sum of (21) to (23)
(22)	Receive compensation or remuneration commensurate with the activities.	1・2・3・4	
(23)	Expenses related to activities are covered by compensation/remuneration for activities and/or grants, etc., and not rely on out-of-pocket expenses.	1・2・3・4	Average

IV VALUE

8 Principle

(24)	The basic principles, mission and ideals of your activity are clarified and can be explained to others.	1・2・3・4	Sum of (24) to (26)
(25)	The nature and characteristics of your arts therapy activities are clarified and can be explained to others.	1・2・3・4	
(26)	Have an attitude of pursuing the essence of art and its expression.	1・2・3・4	Average

Appendix III EAT Rubric

	1	2	3	4
Knowledge	Knowledge is limited or fragmental. Have not systematically studied arts therapy.	Have basic level of knowledge of arts therapy.	Have sufficient knowledge of arts therapy required in your specialized activity area. Able to explain it to others.	Have a systematic and comprehensive knowledge of arts therapy in general as well as familiarity with various disciplines related to your expertise. Able to explain it to others.
Skill	Have very few or a few experiences of EAT practice.	Able to provide EAT based on the methods mastered.	Able to provide safe and efficient EAT with the methods mastered.	Able to provide safe and effective EAT with originally devised methods as well as the methods mastered. Able to customize the methods and process flexibility according to a client's condition.
Psychological Safety	Take no measures for psychological safety in the activities. Do not think of any troubles out of your control that may possibly occur. Have no way to deal with troubles.	Readiness to anticipate possible troubes to some extent. Able to see if the matter is beyond your skill and take some measures.	Aware of the need and importance to secure psychological safety for clients in the activities and necessary measures are taken.	Able to secure psychological safety in the activities and respond to unexpected trouble swiftly and properly. Able to provide proper alternative solutions such as to introduce the client to other relevant agencies, or experts in case that the matter is beyond your skill.
Self-exploration/ Self-understanding	Never had any self-insight about yourself. Your own psychological issues are untouched or unrecognized.	Have some degree of self-insight about yourself. Aware of your psychological issues and work on to find ways to solve them.	Have sufficient self-understanding of your internal and essential propensities. Your own psychological issues are resolved.	Have sufficient and profound understanding of your deep internal and essential propensities. Your own psychological issues have been overcome. Periodically take training analysis or supervision.
Networking/ Cooperation	Have no personal connections or fellows to ask or consult in case of trouble. Have no way of gathering information about activities.	Trying to have reliable personal connections or fellows. Able to get information when you need.	Have reliable personal connections or fellows. Gathering information continuously.	Have reliable connections to contact professionals in the area of your expertise and other related areas. Belong to academic societies. Regularly participate or make presentations in conferences.
Management / Sustainability	Neither think of the sustainability of activities nor the managerial side of your activities for improving sustainability.	Working on or making efforts to raise the sustainability of activities.	Working sustainably regardless of whether paid or not.	Working sustainably and constantly, and have the potential to develop activities.
Principle	The principles and mission are not clarified or have never thought of them.	The principles and mission are clarified and verbalized to some extent.	The principles and mission are clarified and can be explained.	The principles and mission are firmly established and cleary verbalized. Able to deliver the purpose and ideals of your activities to others.

Michi Komura received her MA from Osaka University, Osaka, Japan in 1993. Her professional career includes work experiences as a teacher of English at Sakai Municipal Commercial High School, the deputy secretary general at Osaka NPO Center, a chief producer at Osaka 21st Century Association, and others. She has been engaged in art activities in the community as a board member of a non-profit organization since 2002 and joined the study of Empowerment Arts Therapy in 2012. She has been engaged in Collaborative Online International Learning (COIL) and social innovation education at Osaka Metropolitan University since 2019. She also teaches history of art at the Osaka Institute of Technology for many years. Her current research interests include art activities in the community, empowerment of the non-profit sector, and social innovation.

Midori Ishihara received her Ph.D. from the Graduate School of Letters, Osaka University, Osaka, Japan, in 2002. She has been engaged in the study of Empowerment Arts Therapy since 2008. She is a member of the ATAS Laboratory and a visiting researcher at the Konan Institute of Human Sciences.

Hajime Kaneko received his MA from Osaka City University (changed to Osaka Metropolitan University in 2022), Osaka, Japan, in 1995. He is an Associate Professor at Kobe University of Future Health Sciences. His major is sociology. He was the research representative of the "Arts Therapy Activity Survey in Japan" (2012-2014) and "Requirements and Evaluation Method for 'Empowerment Arts Therapy'" (2015-2018). Both research projects were funded by the Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science.

Akihiro Hayashi received his MBA and Ph.D. from the University of Tsukuba, Tokyo, Japan in 1997 and 2010, and his Ph.D. in Software Engineering from Nanzan University, Nagoya, Japan, in 2019. His professional career includes global companies such as Motorola USA, NTT, and IBM Japan. In 2018 he became a Professor at the Department of Information Design, Faculty of Informatics, Shizuoka Institute of Science and Technology. His current research interests include process improvement, quantitative project management, operating research, and management strategy. He is a member of IEICE and INFSOC.

Industrial Paper

Firmware Distribution with Erasure Coding for IoT Devices

Takenori Sumi[†*], Yukimasa Nagai[†], Jianlin Guo[‡] and Hiroshi Mineno[*]

[†]Information Technology R&D Center, Mitsubishi Electric Corporation, Japan
[*]Graduate School of Science and Technology, Shizuoka University, Japan
[‡] Mitsubishi Electric Research Laboratories, USA
Sumi.Takenori@dc.MitsubishiElectric.co.jp

Abstract – Due to the increase in the number of IoT devices connected to the internet, 920 MHz frequency bands for wireless communication systems are attracting attention for various IoT applications, e.g., environmental monitoring, smart metering, process monitoring & control. With wireless communication systems on 920 MHz having the features of long distance, low rate and low power consumption, a huge number of IoT devices distributed in wide area can be connected to communication networks. When distributing the same data to IoT devices such as firmware distribution during operation, improving the efficiency of distribution method becomes an issue. We propose a new firmware distribution method with erasure coding for IoT devices. Our computer simulation result shows that proposed method improves the efficiency of distribution by 1.7 times for single-hop networks and by 1.5 times for multi-hop networks compared with conventional method and achieves higher spectrum efficiency.

Keywords: IoT, Firmware Distribution, Erasure Code

1 INTRODUCTION

In addition to smart phone, laptop and tablet connecting to the internet, IoT devices such as sensors are getting connect to the internet. Due to the increase in the number of IoT devices connected to the internet, 920 MHz frequency bands for wireless communication systems are attracting attention for various IoT applications, e.g., environmental monitoring application for location, temperature, humidity and water level, smart metering application, process monitoring & control application. Wireless communication systems on 920 MHz (Sub-1 GHz) have the features of long distance, low rate (several 10 kbps – 100 kbps) and low power consumption for conventional standards, e.g., IEEE 802.15.4g [1], LoRaWAN [2], SigFox [3], but the higher data rate up to several Mbps is also considered for applications such as infrastructure monitoring, surveillance camera in IEEE 802.11ah [4][5], marketed as Wi-Fi HaLow. For long life IoT devices, the firmware update for a large number of IoT devices is also considered as new IoT application. Since 920 MHz has the features of long distance from several 100 m to several km, a large number of IoT devices are deployed in the area. Thus, firmware distribution for IoT devices using 920 MHz with narrow band is a challenge. Efficient firmware distribution method for a large number of IoT devices should be considered. In this paper, we focus on 920MHz for IoT applications and propose the new firmware distribution method using erasure coding to achieve higher efficiency for limited radio frequency. Furthermore, since firmware distribution for IoT devices does not require real-time update, the hardware performance such as memory and CPU power for using erasure coding is not a major issue. Generally, updates are expected to occur within tens of minutes to hours for IoT devices.

The rest of this paper is organized as follows. Section II presents related work. Section III describes the proposed firmware distribution with erasure coding. Section IV shows the simulation architecture and results for various conditions. Finally, we conclude our paper in Section V.

2 RELATED WORK

There are existing researches for wireless communications using 920 MHz and firmware distribution. Since 920 MHz is narrow bands compared to 2.4 GHz and 5 GHz for ISM band, special regulation for “10 % transmission duty cycle” and “longer backoff mechanisms” are applied in Japan.

Throughput performance has been demonstrated in [6] and [7], which focus on the PHY and MAC protocol enhancement for higher-throughput, protocol efficiency and delay via simulation and measurement result using prototypes. For example, D. Hotta et al. introduce the performance of multi-hop routing construction using Wi-SUN FAN (Field Area Network) prototypes based on IEEE 802.15.4g FSK PHY [8].

Japanese standard ARIB STA-T108 (20 mW, unlicensed) defines the use of IEEE 802.15.4g system from 920.5 – 928.1MHz (7.6 MHz bandwidth), but the ARIB STA-T107 (250 mW, passive system) and the ARIB STD-T108 (250mW, licensed/registered) also define operation from 920.5 – 923.5 MHz (3.0 MHz). Therefore, 923.5 – 928.1 MHz (4.6 MHz bandwidth) is the only reasonable frequency band for IEEE 802.15.4g applications in the unlicensed spectrum. IEEE 802.15.4g is regulated to operate over 200 kHz bandwidth channel in the Sub-1 GHz band. Even low duty cycle constraint applied in the Sub-1 GHz band, e.g., Japanese and European standard allow up to 10% transmission duty cycle [9]-[12] for the number of IoT Devices increased with various standards. Therefore, ensuring higher efficiency for spectrum use in the Sub-1 GHz is clearly important.

In IoT applications, such as environmental monitoring applications for location, temperature, humidity, and water quantity, as well as smart meter applications, network topology is constructed around sensor information collection nodes. IEEE 802.15.4g / Wi-SUN uses RPL (IPv6 Routing Protocol for Low-Power and Lossy Networks) to construct network topology. In RPL, network topology is constructed by DODAG (Destination Oriented Directed Acyclic Graph). In DODAG, each node only has a route to the root, as shown in Figure 1. Routes are constructed by sending and receiving

DODAG Information Object (DIO) messages to and from neighboring nodes. Thus, each node can know the Rank and other information of its neighbors.

Considering firmware distribution, firmware is sent from Root in Figure 1 to each node. When distributing the same firmware data to a large number of nodes, it is more efficient to distribute the firmware data using multicast instead of sending it to each node via unicast. However, there is an issue of bandwidth congestion if the firmware data is transmitted by flooding, in which all nodes transfer the firmware data. Therefore, Multi-Point Relay (MPR) has been proposed as a more efficient transfer method [14]. According to [14], finding the smallest relay set is NP hard, but a heuristic relay node select is proposed. In the heuristic method, each node selects one-hop neighbors as a relay node to cover all two-hop neighbors. This allows for efficient distribution by limiting the number of nodes to which firmware data is transferred. Furthermore, to improve the reliability of delivery, a method using LT code, one of erasure codes, in MPR has been proposed [15]-[17]. These schemes improve delivery efficiency and reliability by network coding with LT codes on the MPR-constructed transfer path.

However, MPR requires obtaining information on two-hop nodes in addition to one-hop nodes. In RPL, only the information of neighbor nodes is notified, and the information of two-hop nodes is not necessary to notify sensor data to the root. Furthermore, as mentioned earlier, the 920 MHz band is narrow and the 10% duty cycle in Japan and Europe makes it difficult to inform additional information among nodes for firmware distribution. As the propagation environment fluctuates, the connection relationships of the nodes also change, and the connection relationship information must be updated among all nodes.

Therefore, this paper proposes a scheme to determine the forwarding node for firmware distribution without prior downward route construction, and a scheme to improve distribution efficiency with erasure coding.

3 PROPOSED METHOD

3.1 Challenges in Firmware Distribution

Firmware distribution to IoT devices requires sending the same firmware data to many IoT devices via wireless communication. Firmware distribution using unicast is inefficient because the same firmware data is sent to each device individually. Therefore, it is considered more efficient to distribute to a large number of IoT devices using broadcast. However, it is difficult to deliver all firmware data to all IoT devices in a single communication because wireless communication suffers from packet errors due to low received signal power and/or interference from other wireless systems. As shown in Figure 2, packet errors occur randomly in IoT devices, requiring large number of packets to be retransmitted. Therefore, even when broadcasts are used for firmware distribution, there are issues with distribution efficiency. In this paper, we propose a method to reduce the number of transmitted packets in firmware distribution by using erasure coding.

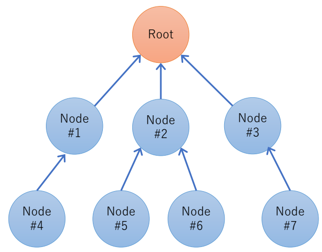


Figure 1: Routing Construction using RPL

3.2 Erasure Code

In this paper, RC QC-LDPC (Rate Compatible Quasi-Cyclic-LDPC) [18] is used as an erasure code for firmware distribution. In general, erasure encoding are used for distributed storage. In distributed storage, it is sufficient to be able to recover the loss of some data out of N, e.g., when some of the multiple storage devices fails. However, in wireless communication, it is necessary to cope with packet errors for some long time in the time axis direction because of bursty packet errors. RC QC-LDPC can handle bursty packet errors because it can take a large number of information packets and redundant packets, as described below. In the related studies described in Chapter 2, network coding was applied using LT code based on the forwarding path constructed by MPR. In this paper, however, standard erasure coding is used instead of network coding because the downward forwarding path is not constructed in advance. As shown in Figure 3, RC QC-LDPC generates redundant packets by erasure coding information packets that divide the data to be transmitted into packets of a certain data length. The data length of information packets and redundant packets are the same. The number of information packets, K, is a multiple of 36, and the number of redundant packets, M, is equal to K or twice as large. RC QC-LDPC can attempt to decode the transmitted data if the total number of received information packets and redundant packets is K or more. Even if some of the received information packets are missing, the transmitted data can be recovered using the redundant packets. Figure 4 shows the relationship between "coding redundancy rate" and decoding success rate. Here, the "coding redundancy rate" is defined as the sum of the number of information packets K_{rx} and the number of redundant packets M_{rx} at decoding and the original number of information packets K, using the following formula (1).

$$Coding\ Redundancy\ Rate = \frac{K_{rx}+M_{rx}}{K} \tag{1}$$

Figure 4 shows that RC QC-LDPC does not ensure successful decoding even when the redundancy is greater than 1. The higher the number of information packets K, the higher the decoding success rate for the coding redundancy rate. However, the larger the number of information packets K, the more memory and CPU resources are required for the decoding process.

3.3 Firmware Distribution with Erasure Code

Before describing firmware distribution in multi-hop networks, we describe how loss-correcting codes can be applied to firmware distribution in single-hop networks. In the proposed firmware distribution method, as shown in Figure 5, the transmitter first erasure codes the firmware data (K information packets in the figure) to generate M redundancy packets ((1) in the figure). If firmware data is large, the transmitter performs erasure coding multiple times.

Then, the Transmitter sends up to a total of K information packets and redundant packets ((2) in the figure). Figure 7 shows the format of information packets and redundant packets. The Information packet and the redundant packet contain packet type indicating whether it is an information packet or redundant packet, sequence number, packet index, and payload. The sequence number is incremented each time the transmitting station encodes firmware data and is used to identify which data the response is to. The packet index indicates the index of information packets or redundant packets. After K packets have been sent, depending on the number of packet errors in the IoT device, additional information packets and redundant packets are sent until the IoT device can decode the firmware data ((3) in the figure). Here, the additional packets to be sent are those not sent in (2) in the figure. The IoT device uses the received K or more packets to decode and recover the firmware data ((4) in the figure). As described in section 3.2, RC QC-LDPC may fail in decoding. If decoding fails, additional packets are sent from the transmitter and the IoT device performs the decoding process again.

To implement the proposed method, IoT devices need to inform the transmitter which packets are missing. In the proposed scheme, IoT devices notify the transmitter of a NACK (Negative Acknowledgment) packet and an ACK (Acknowledgment) packet in the sequence shown in Figure 6. The NACK packet is used to notify which packets are missing, the ACK packet is used to notify that decoding was successful. The packet formats of NACK and ACK packets are shown in Figure 8. Multicast UDP packets is used to send NACK and ACK packets, and Figure 8 shows only the UDP payload. The NACK packet contains packet type indicating whether it is NACK or ACK, destination node identifier, sequence number and a bitmap indicating packet loss. The length of the bitmap is the sum of the number of information packets K and the number of redundant packets M. The ACK packet contains packet type, destination node identifier and sequence number.

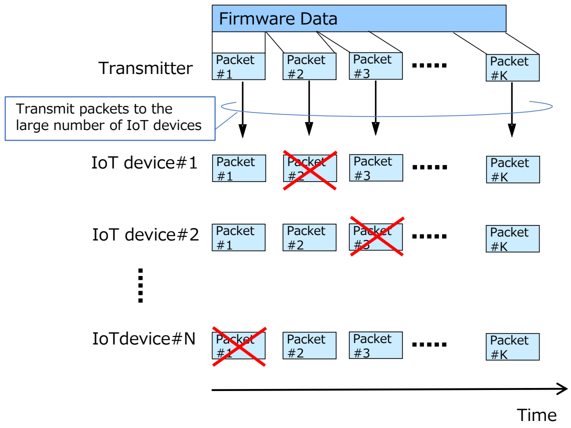


Figure 2: Firmware Distribution for IoT Devices

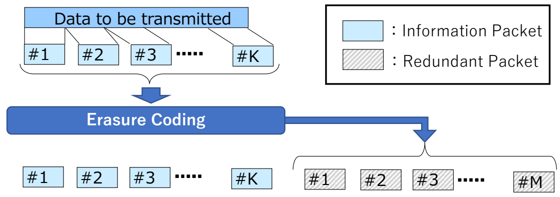


Figure 3: Erasure Coding

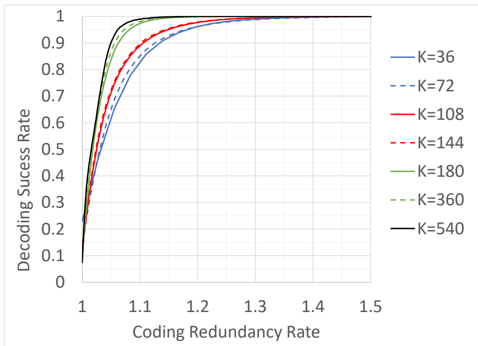


Figure 4: Redundancy Rate vs. Decoding Success Rate

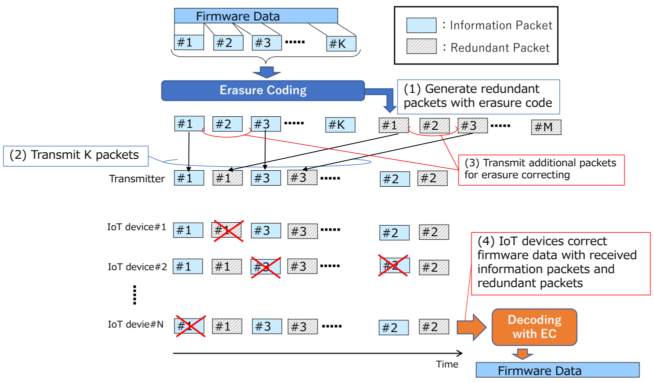


Figure 5: Proposed firmware distribution with EC

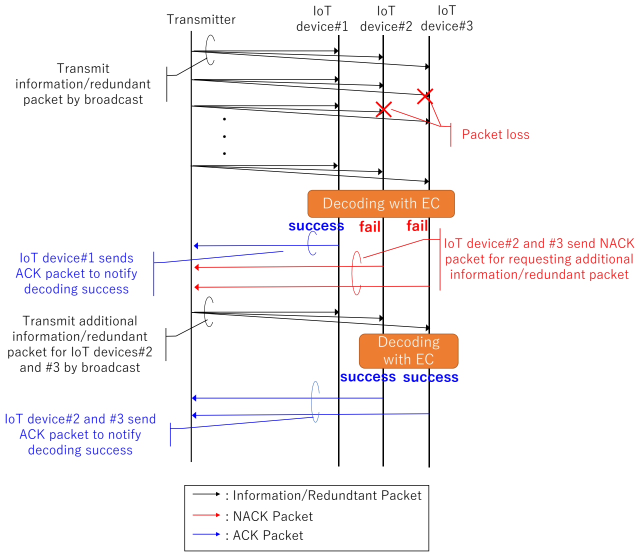


Figure 6: Packet sequence of the proposed method

1 Byte	1 Byte	2 Byte	2 Byte	
Packet type	Hop count	Sequence number	Packet index	Payload

Figure 7: Packet format of information packet and redundant packet

NACK Packet	1 Byte	2 Byte	2 Byte	(K + M)/8 Byte
	Packet type	Destination	Sequence number	Packet loss bitmap 1: Received, 0: Packet lost

ACK Packet	1 Byte	2 Byte	2 Byte
	Packet type	Destination	Sequence number

Figure 8: Packet formats of NACK and ACK packet

3.4 Firmware Distribution in Multi-hop Networks

Section 3.3 described the firmware distribution method in single-hop networks. In this section, we describe the firmware distribution method in multi-hop networks.

As discussed in Chapter 2, in IoT applications, each node only needs to know the path to the root node in order to collect sensor data to the root node. However, firmware distribution is distributed from the root node via a downward path. To improve the efficiency of distribution, the number of forwarding nodes needs to be reduced, as in MPR. The proposed method reduces the number of forwarding nodes by using ACK or NACK packets used to confirm the delivery of firmware distribution. In the multi-hop network shown in Figure 9, firmware data can be delivered to all nodes if only Node#1 and Node#3 perform forwarding. The red dotted lines in the figure indicate possible communication paths. In the proposed scheme, firmware data is divided into packets of a certain size and transmitted with erasure coding for each K packets. After receiving the first K packets, each node sends ACK or NACK to the source node by multicast. Here, if Node#4 sends an ACK or NACK packet to Node#1 first, Node#5 can receive the ACK or NACK packet sent by Node#4. If Node#5 has not yet sent an ACK or NACK packet, Node#5 can make Node#1 the destination for ACK or NACK packets, so that the destination for ACK or NACK notification is the same as Node#4. After each node sends its first K packet, only the node that receives an ACK or NACK packet from the next hop node will perform the transfer, thereby reducing the number of nodes transferring firmware data in a multi-hop network and improving delivery efficiency. Furthermore, the number of hops in the information packet or redundant packet of each node and the reception power of the packet can be used to determine the candidate destination for ACK or NACK packet notification, so that the network topology can be maintained and a node with stable reception power can be selected as the ACK or NACK packet destination. In addition, the timing of ACK or NACK notification is shifted for each node by a random number to consolidate the ACK or NACK notification destinations.

4 SIMULATION EVALUATION

We evaluate the proposed method using computer simulation. Table 1 shows simulation parameters. In the computer simulation, IEEE 802.15.4g, which is used in IoT devices in the 920 MHz band, was used as the wireless communication method and evaluated in an environment where IoT devices are connected in a star network or mesh network from a transmitter. IEEE 802.15.4g is a communication method used for smart utility networks [19], and according to TTC JJ-300.10 [20], it is also used for smart meters in Japan. In this simulation, the parameter values of MAC and PHY of IEEE 802.15.4g and JJ-300.10 are used. Assuming environmental monitoring using smart sensors, the number of IoT devices is assumed to range from 1 to 100 in consideration of installation on office (indoor use case) and university campus (outdoor use case) for single-hop network, and from 100 to 1000 in consideration of installation on suburban area for multi-hop network. The firmware data is divided into 248 Bytes each and multicast to the IoT device as the payload of information packets. RC QC-LDPC was used as erasure code, and the number of information packets K and redundancy packets M were set to 36 - 540. In the computer simulation, the conventional firmware distribution without erasure code also broadcasts the firmware data to the IoT devices by dividing it into 248 Bytes each, as in the proposed method. Then, after the K-packet transmission is finished, the IoT devices send ACK or NACK to the transmitter to notify whether it needs to retransmit data. This is to ensure that the opportunities to send ACK and NACK packets are the same for the conventional and proposed methods. In the conventional method, the number of redundancy packets M is 0 because erasure coding is not performed for the firmware data.

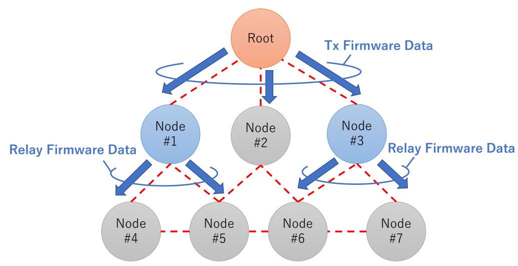


Figure 9: Firmware distribution in multi-hop networks

Table 1: Simulation parameters

Firmware distribution parameters	Value
The number of IoT devices	Single-hop: 1 – 100 Multi-hop : 100 - 1000
Erasure code	Rate-Compatible QC-LDPC
K, the number of information packets	36, 72, 108, 144, 180, 360, 540
M, the number of redundant packets	Conventional method: 0 Proposed method: same as K
Firmware distribution packet length	248 Byte
Firmware distribution ACK packet	1 Byte
Firmware distribution NACK packet	1+(K+M)/8 Byte
6LowPAN and UDP header	53 Byte
PHY/MAC	IEEE 802.15.4g
MAC parameters	Value
macMinBE	5
macMaxBE	8
LIFS	1000 us
aUnitBackoffPeriod	1130 us
phyCcaDuration	130 us
aTurnaroundTime	1000 us
tack	1000 us
MAC header size	11 Byte
FCS	2 Byte
PHY parameters	Value
Data rate	100 kbps
Modulation	2-FSK
Modulation index	1.0
Frequency	923.7 MHz
Channel spacing	400 kHz
Propagation Model	SEAMCAT Extended Hata Model (Suburban)
Antenna height	1.5 m

4.1 Single-hop Network Evaluation

First, the relationship between the number of information packets K and effective throughput is shown in Figure 10, which shows the simulation results when firmware is distributed to 20 IoT devices and PER is 10%. Since a PER of 10% or less is often set for wireless communication systems to take operations into account [19], the simulations for single-hop network in this paper are based on an evaluation at a PER of 10%. According to the figure, the throughput of the proposed method with erasure coding is higher than that of the conventional method without erasure coding, regardless of the value of K. As the value of K increases, the effective throughput of the proposed method increases. As mentioned in section 3.2, this is because the larger K is, the higher decoding success rate at the same redundancy rate. However, the memory and decoding processing load increases as K increases. The relationship between K ant the memory size required for erasure coding is shown in Figure 11. The required memory size increases linearly with K. On the other hand, Figure 10 shows that throughput improvement slows down when K is greater than 180. Furthermore, if K is greater than 180, the required memory size exceeds 100 KB. Therefore, for embedded devices with limited memory size, it is better to set K to 180. The subsequent simulation evaluation will be performed for the case where the number of information packets K is 180, where the effective throughput increase rate by the proposed method is relatively high.

Next, Figure 12 shows the simulation result on the relationship between PER and effective throughput when firmware is distributed to 20 IoT devices with the number of information packets K set to 180. The figure shows that if the PER is 1% or higher, the effective throughput is higher with the proposed method than with the conventional method. However, the effective throughput at a PER of 0% with no packet errors is approximately 40 kbps for the conventional method, but 33 kbps for the proposed method, which is lower than the conventional method. As mentioned in Section 3.2, this is because RC QC-LDPC codes may not succeed in decoding even if a total of K or more information packets and redundancy packets are received. Additional packets need to be sent from the transmitter to the IoT devices. Since packet errors occur in wireless communication systems, the proposed method can be applied to improve the effective throughput in an environment with a PER < 10%, where wireless communication systems are normally operated.

Finally, the relationship between the number of IoT devices and effective throughput is shown in Figure 13, which shows simulation results when the number of information packets K is 180 and PER is 10%. The figure shows that the effective throughput of the proposed method is higher than that of the conventional method when firmware distribution is performed to two or more IoT devices. In the case of a single IoT device, the effective throughput is lower than the conventional method due to the significant impact of the possibility of unsuccessful decoding even if the number of received packets is K or more. For 20 IoT devices, it is 1.60 times the effective throughput of the conventional method; for 50 devices, it is 1.73 times; and for 100 devices, it is 1.76 times. However, as the number of IoT devices increases, the effective throughput decreases: 29.7 kbps for 20 IoT devices, 24.8 kbps for 50 devices, and 19.8 kbps for 100 devices. This is likely due to an increase in NACK packets.

The transmission duty cycle of the transmitter in the simulation of Figure 13 is shown in Figure 14. As mentioned in Chapter 2, Japanese and European standard allow up to 10% transmission duty cycle. Figure 14 shows that regardless of the number of IoT devices, the transmission duty cycle of the transmitter exceeds 10% for both the proposed and conventional methods. The transmission duty cycle of the proposed method is lower than that of the conventional method, and it tends to be lower when the number of IoT devices is larger. Therefore, in Japan and Europe, when the transmission duty cycle is 10% or less, the difference between the effective throughput of the proposed method and the conventional method is expected to widen. The firmware delivery time to 100 IoT devices is shown in Figure 15 considering that IoT devices are communicating in normal operation, i.e., notifying sensor data, the firmware distribution time was evaluated with the firmware delivery bandwidth usage rate as 1%, 5%, and 10%. The difference in firmware delivery time when erasure coding is applied or not applied is about 8.5 hours when the firmware data size is 256 KB, 1.7 hours when the bandwidth utilization is 1%, 1.7 hours when 5%, and 0.85 hours when 10%. However, when the firmware data size increases, for example, to 1024 KB, the difference is 34 hours at 1%, 13.6 hours at 5%, and 6.8 hours at 10%. If the normal state communication of the IoT device is bandwidth-hungry and the firmware data size is 1024 MB or larger, the proposed method can improve the firmware delivery time by half a day to a day or more.

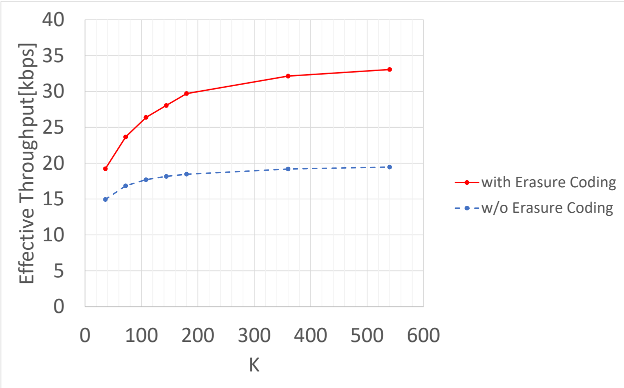


Figure 10: K vs. Throughput (PER: 10%, 20 devices)

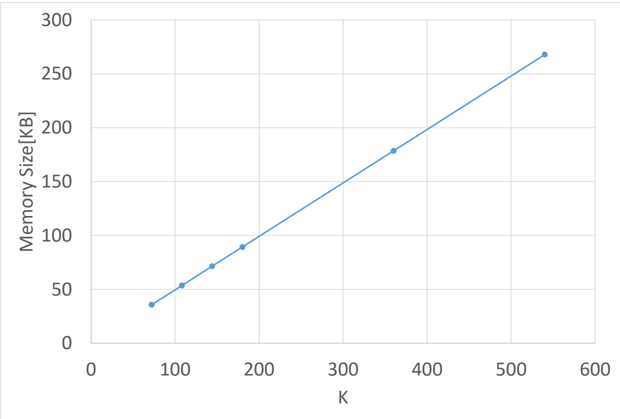


Figure 11: K vs. Memory Size

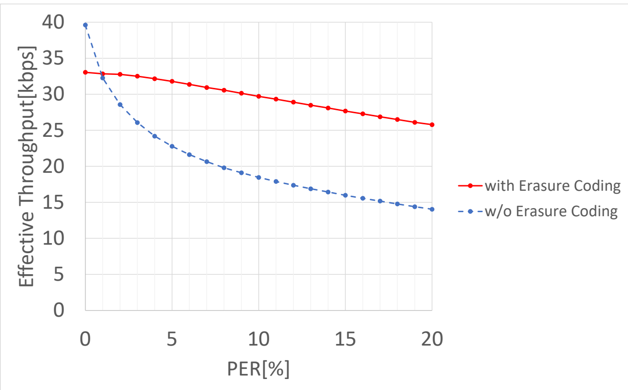


Figure 12: PER vs. Throughput (K: 180, 20 devices)

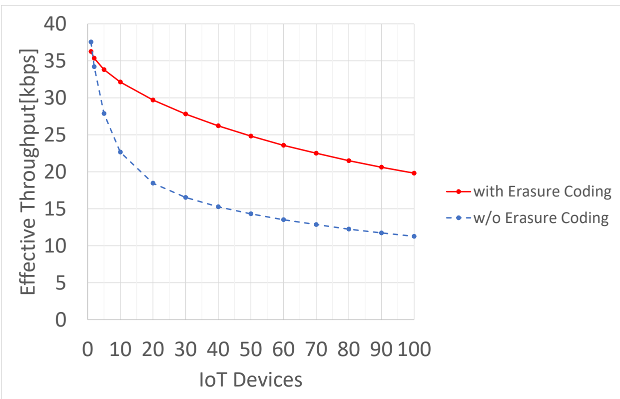


Figure 13: The number of IoT devices vs. Throughput (K: 180, PER: 10%)

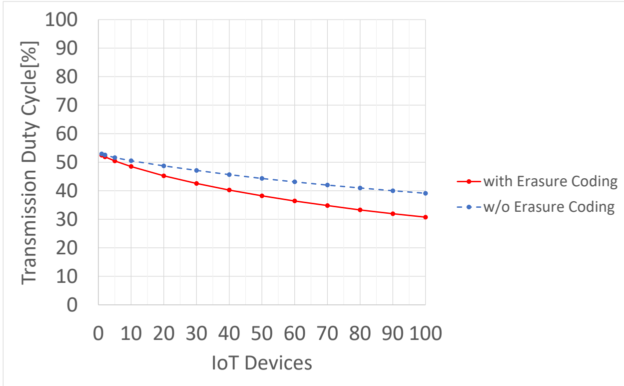


Figure 14: The number of IoT devices vs. Duty Cycle (K: 180, PER: 10%)

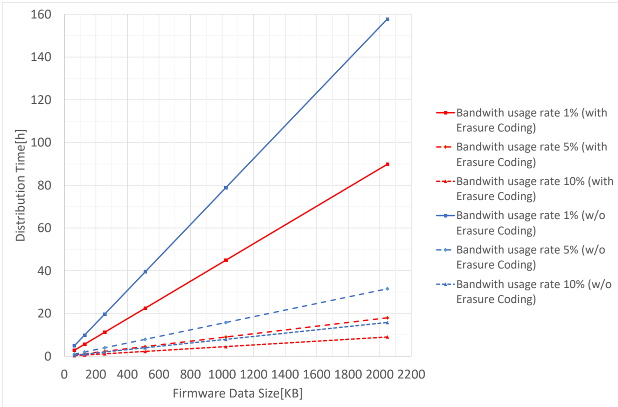


Figure 15: Firmware Data Size vs. Distribution Time (K: 180, PER: 10%, 100 devices)

4.2 Multi-hop Network Evaluation

For the evaluation of the multi-hop network, firmware data transmitter is placed at the center of a circle with a radius of 500 m, and 100 to 1000 nodes are randomly placed within the circle. Suburban of SEAMCAT Extended Hata Model was used as the propagation model and the received power was calculated. Figure 16 shows the relationship between distance and received power. With 2-FSK, the modulation scheme used in the simulation, reception is stable down to a minimum of -90 dBm, which is about 89 m from the figure. In MPR, 1-hop and 2-hop nodes were calculated with an upper limit of 80 m distance to account for packet errors due to interference. For the number of information packets K of erasure coding, 180 was used based on the single-hop evaluation. In the single-hop evaluation, since only the transmitter sends firmware data packets, frame collisions did not occur, so the packet error rate was set and packet errors were generated randomly, but in multi-hop, multiple nodes transfer firmware data packets, so frame collisions occur. Therefore, the packet error rate was not set, and the evaluation was based on packet loss due to frame collisions.

First, Figure 17 shows an example of terminal placement under the MPR method and the proposed method in which the relay node is determined by ACK or NACK packets for firmware distribution. The figure shows an example with 500 nodes. The black dotted line connecting each node indicates that firmware data is being transferred by relay. In the proposed method, 244 out of 500 nodes are set as relay nodes, while in MPR, 237 nodes are set as relay nodes, which are almost equal in number. It can be seen that the proposed method can set up the same number of relay nodes as MPR without having any routing information other than the uplink route by RPL.

Next, Effective throughput is shown in Figure 18 for the proposed method, MPR, and flooding with and without Erasure coding, respectively. The figure shows that the proposed method and MPR achieve almost the same effective throughput. Compared to the flooding case, the proposed method improves effective throughput by about 20-25%. In addition, the effective throughput is improved by 52-57% when Erasure coding is used in the proposed method compared to the case where Erasure coding is not used in flooding. The average and maximum number of hops are shown in Figure 19. The figure shows that for 400 or more IoT devices, the average and maximum hop counts are almost equal. This means that the decrease in effective throughput is not caused by an increase in the number of hops, but by an increase in the number of terminals.

Finally, Figure 20 shows the packet transmission redundancy. The redundancy of packet transmission is expressed by the following formula (2).

$$Packet\ Transmission\ Redundancy = \frac{\sum_{n=1}^{N} P_{FWn}+P_{ACKn}+P_{NACKn}}{P_{FW}KN} \quad (2)$$

Here, P_{FWn} is the total size of firmware data packets (information packets or redundant packets) sent by node#n, P_{ACKn} is the total size of ACK packets sent by node#n, P_{NACKn} is the total size of NACK packets sent by node#n, P_{FW} is the size of a firmware data packet, K is the number of information packets for erasure coding, and N is the number of nodes. In other words, the redundancy of packet transmission is calculated based on the size of the transmitted data when all nodes except the firmware data transmission node transmit the firmware data once. The figure shows that the proposed method and MPR have almost the same level of redundancy, ranging from 0.5 to 0.6. In the case of flooding, the redundancy is about 1.1 to 1.2, indicating that the proposed method can deliver firmware data in about half the transmission time compared to flooding.

5 CONCLUSION

We proposed the new firmware distribution method using erasure coding to achieve higher efficiency for limited radio frequency and the method for selecting relay nodes in a multi-hop network without additional downward routing information. The performance of the proposed method was evaluated by computer simulation and compared to the effective throughput of conventional firmware distribution methods. For single-hop network, by applying the proposed method to firmware distribution, the effective throughput was found to be 1.60 times higher than that with the conventional method when there are 20 IoT devices receiving firmware data, 1.73 times higher when there are 50 devices, and 1.76 times higher when there are 100 devices. For multi-hop network, by applying the proposed method to firmware distribution, effective throughput was found to be about 1.5 times. The memory size required for encoding is about 100 KB for the erasure encoding scheme used in this paper, which is considered to be operable on mid- to high-performance models of embedded CPUs. However, we plan to investigate memory-saving erasure coding schemes that can operate on even lower-performance models in the future. Furthermore, in this paper, we evaluated the performance with computer simulation. Our future work is conduct evaluations using actual equipment taking into account the amount of memory and other factors.

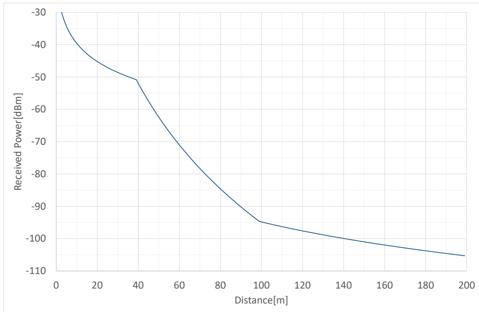


Figure 16: Distance vs. Received power

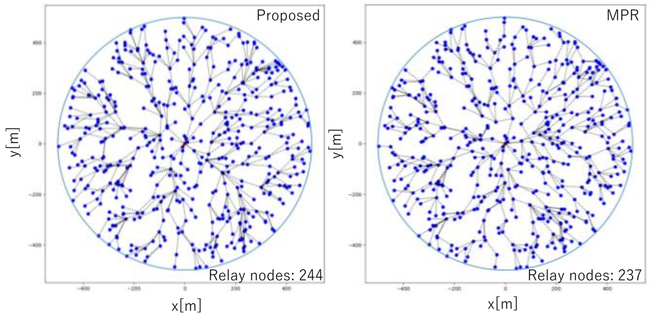


Figure 17: Example of relay node selection (500 nodes)

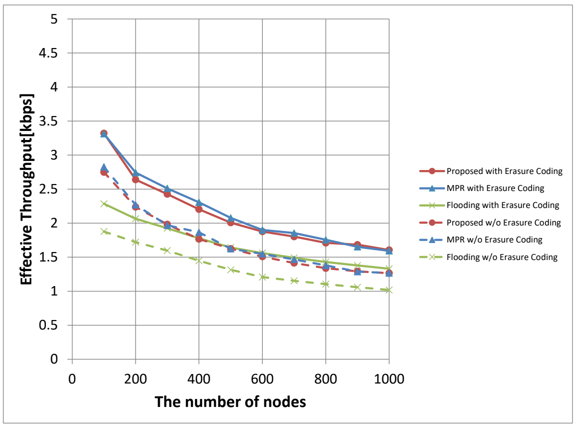


Figure 18: Effective throughput (Multi-hop)

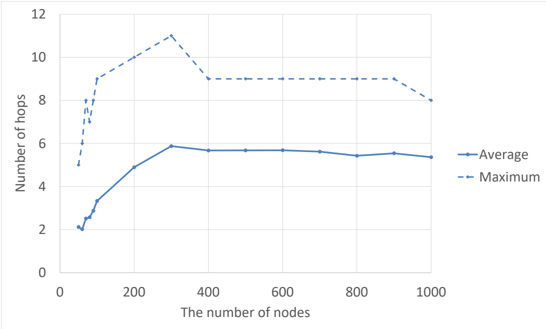


Figure 19: Average and maximum hops

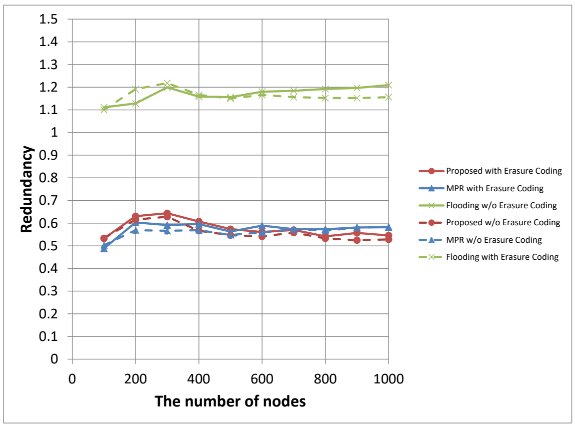


Figure 20: Packet transmission redundancy (Multi-hop)

REFERENCES

[1] IEEE 802.15.4g-2012, "IEEE Standard for Local and metropolitan area networks – Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs) Amendment 3: Physical Layer (PHY) Specifications for Low-Data-Rate, Wireless, Smart Metering Utility Networks," IEEE 802.15.4g-2012, IEEE, (2012).

[2] LoRa Alliance, "LoRaWAN 1.1 Specification," (2017).

[3] Sigfox, "Sigfox connected objects: Radio specifications Rev. 1.6," (2022).

[4] IEEE 802.11ah-2016, "IEEE Standard for Information technology-Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 2: Sub 1 GHz License Exempt Operation," IEEE 802.11ah-2016, IEEE, (2016).

[5] 802.11ah Promotion Councill, "IEEE802.11ah Use case," https://www.11ahpc.org/pdf/usecase-eng.pdf, (2019).

[6] C.-S. Sum, F. Kojima, and H. Harada, ''Performance analysis of a multiPHY coexistence mechanism for IEEE 802.15.4g FSK network,'' in Proc. IEEE Wireless Communication Network. Conf. (WCNC), pp. 41–46, (2013).

[7] F. Righetti, C. Vallati, D. Comola, and G. Anastasi, ''Performance measurements of IEEE 802.15.4g wireless networks,'' in Proc. IEEE 20th Int. Symposium World Wireless, Mobile Multimedia Network (WoWMoM), pp. 1–6, (2019).

[8] D. Hotta, R. Okumura, K. Mizutani, and H. Harada, ''Stabilization of multi-hop routing construction in Wi-SUN FAN systems,'' in Proc. IEEE 17th Annual Consumer Communication and Network Conference (CCNC), pp. 1–6, (2020).

[9] ARIB STD-T107 version 1.1, "920MHz-Band RFID Equipment for Specified Low Power Radio Station," (2017).

[10] ARIB STD-T108 version 1.4, "920MHz-Band Telemeter, Telecontrol and Data Transmission Radio Equipment," (2021).

[11] Y. Nagai, J. Guo, P. Orlik, B. A. Rolfe, and T. Sumi, "Proposal for Section 6.1 Japan on Recommended Practice for Local and Metropolitan Area Networks - Part 19: Coexistence Methods for 802.11 and 802.15.4 based systems operating in the Sub-1 GHz Frequency Bands," IEEE 802.19-19/0049r0, https://mentor.ieee.org/802.19/2019/, IEEE 802.19, (2019).

[12] T. Sumi, Y. Nagai, J. Guo, and P. Orlik, "920 MHz Status Update in Japan," IEEE 802.19-18/0084r0, https://mentor.ieee.org/802.19/2018/, IEEE 802.19, (2018).

[13] RFC 6550, "RPL: IPv6 Routing Protocol for Low-Power and Lossy Networks," IETF, (2012).

[14] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint Relaying: An Efficient Technique for Flooding in Mobile Wireless Networks," in The 35th Annual Hawaii International Conference on System Sciences, (2001).

[15] N. Kadi and K. A. Agha, "Optimized MPR-based flooding in wireless ad hoc network using network coding," in 1st IFIP Wireless Days, (2008).

[16] K. A. Agha, N. Kadi, and I. Stojmenovic, "Fountain Codes with XOR of Encoded Packets for Broadcasting and source independent backbone in Multi-hop Networks using Network Coding," in IEEE VTC Spring 2009, (2009).

[17] N. Kadi and K. A. Agha, "MPR-based flooding with distributed fountain network coding," in 2010 The 9th IEIP Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net), (2010).

[18] W. Matsumoto, R. Sakai, H. Yoshida, "Rate Compatible QC-LDPC codes," SITA 2006, pp.387-390, (2006).

[19] IEEE 802.15.4-2020, "IEEE Standard for Low - RateWireless Networks," IEEE 802.15.4-2020, IEEE, (2020).

[20] TTC Standard JJ-300.10 Version 2.2, "Home Network Communication Interface for ECHONET Lite (IEEE802.15.4/4g/4e 920 MHz-band Wireless)," TTC, Version 2.2, (2015).

(Received: March 1, 2023)
(Accepted: June 29, 2023)

Takenori Sumi received B.E. degree and M.E. degree in Faculty of Engineering of Doshisha University, Kyoto, Japan in 2005 and 2007, respectively. He is currently a Ph.D student in Graduate School of Science and Technology, Shizuoka University, Japan. He has been working as a senior researcher at Information Technology R & D Center of Mitsubishi Electric Corporation from 2007. His current research interests include spectrum sharing, wireless coexistence and V2X. He is also engaged in the Wi-Fi architect and network architect for embedded system. IEEE Senior member, IEICE member and IPSJ member.

Yukimasa Nagai received B.E. and M.E. degrees from The University of Electro-Communications, Japan in 1998 and 2000, respectively. In 2022, he received his Ph.D. degree in informatics from Graduate School of Science and Technology, Shizuoka University, Japan. He has been working at Mitsubishi Electric Corporation in Japan (2000-2016, 2020-) and Mitsubishi Electric Research Laboratories in USA (2016-2020). His research work has covered various topics on communication systems, IoT, CN and he has acted as key parts of getting standardization groups formed in Wi-Fi Alliance and IEEE 802.11/15/19. Wi-Fi Alliance Automotive MSTG Vice Chair, IEEE Senior member, and IPSJ member.

Jianlin Guo is a Senior Principal Research Scientist at Mitsubishi Electric Research Laboratories in Cambridge, Massachusetts, USA. He received his Ph.D. in 1995 from University of Windsor, Canada. His research interests include coexistence of the heterogeneous wireless networks, routing and resource management in wireless IoT networks, distributed machine learning, time sensitive networking, edge computing for vehicular networks, networked control systems, and fault detection in industrial automation. He served as the Technical Editor for IEEE 802.19.3 Task Group. He is IEEE Senior Member.

Hiroshi Mineno received his B.E. and M.E. degrees from Shizuoka University, Japan in 1997 and 1999, respectively. In 2006, he received his Ph.D. degree in information science and electrical engineering from Kyushu University, Japan. Between 1999 and 2002, he was a researcher in the NTT Service Integration Laboratories. In 2002, he joined the Department of Computer Science of Shizuoka University as an Assistant Professor. He is currently a Professor. His research interests include Intelligent IoT systems as well as heterogeneous network convergence. He is a senior member of IEEE, IEICE and IPSJ, a member of ACM and the Informatics Society.

Regular Paper

Utility of Remote Workers' Video of a Surrogate Telepresence Robot for Predicting Its Motion[1]

Kosuke Sasaki†, Zijie Yuan†and Tomoo Inoue‡

†Graduate School of Library, Information and Media Studies, University of Tsukuba, Japan
‡Institute of Library, Information and Media Science of University of Tsukuba, Japan
{ksasaki, inoue}@slis.tsukuba.ac.jp

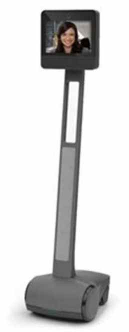


Figure 1: Beam® from Suitable Technologies

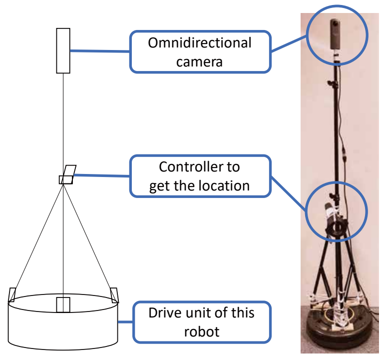


Figure 2: The telepresence robot used in this study

Abstract - This study focuses on environments where people and robots coexist and investigates how to support the coexistence of people and robots. This study assumes a situation in which local workers and telepresence surrogate robots operated from remote locations work in the same office. To investigate how to support the situation, we examined whether a person and a robot can pass by each other by using a technology that projects an image of the person who operates the robot onto the robot, which has been proposed to support physical cooperative work. The experiment results showed that a person could predict the robot's direction of travel and pass by looking mainly at the upper body of the remote person. This suggests the effect of projecting a human figure onto a robot in an environment where humans and robots coexist.

Keywords: Cooperative work, Physical interaction, Nonverbal cue, Telepresence robot

1 INTRODUCTION

Remote work has recently become increasingly popular. Nevertheless, face-to-face work remains essential. This study considers a work environment in which telework and face-to-face work are combined. That is, some workers work in an office while others work at a remote location. In this environment, the remote worker operates robots in the office to communicate with other workers.

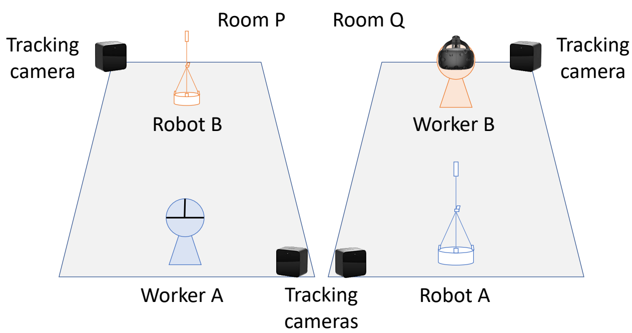


Figure 3: An environment of passing each other

The robot assumed for use in this study is a surrogate telepresence robot, such as the Beam Pro from Suitable Technologies (now Blue Ocean Robotics of Denmark), with a mobile drive unit and a camera or monitor attached to the top of the robot. By using such a telepresence robot, workers in an office can move with the robot and communicate with remote workers via the robot. Such telepresence robots have been used in office environments [2], [3] and in non-office environments [4], [5], and various studies have argued for the usefulness of telepresence robots.

However, previous studies have not paid much attention to the coexistence between humans and robots. There needs to be more understanding of what kind of support is possible when humans and robots freely move through the same space. This study aims to achieve smooth coexistence between humans and robots. In order to examine what kind of assistance is available, this study focuses on a situation in which a person and a robot pass by each other in a narrow space as one

[1]This paper is an extended version of the IWIN proceeding [1].
[2]Home - Beam, `https://suitabletech.com/home` (Visited on Jul 21, 2023)
[3]Blue Ocean Robotics - We Create and Commercialize Robots, `https://www.blue-ocean-robotics.com/` (Visited on Jul 21, 2023)

of the situations in which a person and a robot coexist.

In an environment in which humans and robots coexist, humans and robots must behave cooperatively. Many studies have been conducted to support human-robot coexistence [3], [6]–[8], which is the aim of this research. We must determine what kind of assistance is needed for people and robots to coexist.

In cooperative work in a face-to-face environment or a remote environment with fixed displays, it is known that workers can guess their partners' actions by observing their facial direction, body movements, and other information, thereby being able to work without conflict or collision [9]–[11]. Additionally, whole-body movements are known to provide the viewer with emotional information such as credibility and persuasiveness, which is essential to work collaboratively [12], [13]. However, since robots lack such nonverbal information, humans and robots may not pass by each other well.

Inoue and Yuan have proposed a virtual reality (VR) system for collaborative work with physical movement. In the system, a human and a telepresence robot (as shown in Fig. 2) operated by another person in a remote place passed by each other, which is a form of collaborative work. This system shows the feasibility of a human and a telepresence robot passing by each other in the same workplace [14].

In this study, we examined the method by which Inoue and Yuan initially proposed that a video of a robot operator's whole body in a remote place is shown to a local worker who passes by a telepresence robot. We verified the effect of this method in an environment where humans and robots coexist based on objective and subjective data.

We examined the method by a passing-by task. A worker and a telepresence robot were placed in the same workspace in the passing-by task. Another worker in another workspace operated the telepresence robot. In the experiment, we arranged two pairs of a human and a robot in two workspaces (as shown in Fig. 3). Each worker faced a telepresence robot that moved synchronously with a remote worker and wore a head-mounted display (HMD) that displayed video captured by the camera attached to the robot, which was in a remote place. The worker could walk and see the remote worker's whole body through the HMD.

We analyzed how a worker moved while seeing the remote worker's whole body during the passing-by task based on recorded video of the experiment, a questionnaire survey, and interviews with the workers. The analysis result shows that a worker passing by a telepresence robot tended to look at the upper body of the remote worker's image. Additionally, workers using the system could predict where the robot was moving to [1]. This result suggests that the system displaying the remote worker's image, especially the worker's upper body, onto the robot can support situations in which humans and robots coexist.

2 RELATED WORK

We investigate a situation in which a human and a telepresence robot pass by each other, regarded as remote cooperative work between a local worker and a remote worker. In this chapter, we introduce remote cooperative work support.

2.1 Remote Communication Support Using Flat Fixed Displays

Unlike face-to-face environments, remote environments lack visual and nonverbal information, such as gaze information and body movements. Since nonverbal information is essential for smooth communication, studies have been conducted on communicating visual and nonverbal information with each other to facilitate remote collaboration.

Fixed displays are often used to communicate in remote workspaces. Previous studies have proposed systems that use worker gaze information [15], or that project a shared workspace or background to give users immersion and the feeling of being in the same room [16], [17]. Some proposed systems show images of remote workers on fixed displays. Methods of overlaying remote persons on the background of a fixed display [18] or showing a worker's life-size image on display [19] enable communication as if the workers are in the same space.

Other studies have proposed methods of conveying nonverbal information by projecting images of a remote worker instead of using flat displays [20], [21]. However, methods using flat displays or projections greatly restrict workers' body movements, and these methods cannot support remote cooperative work involving spatial movement.

2.2 Remote Communication Support Using AR and VR Technologies

Systems using VR and augmented reality (AR) technologies with HMDs have also been proposed to solve the problem of movement restriction, including systems that immerse a worker in the same virtual space with a remote worker [22], [23], systems that immerse a worker by displaying the remote workspace [24], and systems that can switch the image displayed on the HMD to other images captured from other viewpoint, such as another worker's point of view or a bird's-eye view [25]. Furthermore, some systems use surrogate avatars of workers for smooth communication [26]–[29]. These previous studies show that AR and VR technologies can reduce restrictions on worker movement. However, in collaborative work in a virtual space, workers do not have physical bodies and cannot interact in real space. Furthermore, using avatars can eliminate the nonverbal information of workers.

2.3 Remote Communication Support Using Telepresence Robots

Several studies have investigated telepresence robots, as we focus on in this study, to support communication in the office. A study by Shen et al. investigated the interpersonal distance between a person and a telepresence robot. American and Indian participants were tested in a situation in which the participant operated a telepresence robot to approach a co-worker working in an office. The results revealed that the worker operating the robot behaved similarly in terms of culture-specific distancing as in a face-to-face environment [2]. Myodo et al. investigated whether telepresence robots could facilitate informal communication in an office setting. Experiments using a telepresence robot to communicate between a remote supervisor and a subordinate in the office revealed that the supervisor's facial expressions were conveyed more strongly when the telepresence robot was used [3].

There have been studies using telepresence robots in environments other than the office. Newhart and Olson's study investigated the use of telepresence robots by children who cannot attend school due to illness. While they showed that the use of telepresence robots in schools could provide learning opportunities for children who are unable to attend school, they also pointed out that there is room for improvement in the school environment where telepresence robots are used, such as physical barriers in the school and the operability of the robots [4]. Another study examined the use of telepresence robots to allow remote users to attend academic conferences. In this study, telepresence robots were arranged at ACM international conferences such as CHI, CSCW, and UBICOMP/ISWC to observe the reactions of local attendees and their interactions with the telepresence robots' in conference activities. As a result, they revealed several insights about the requirements for telepresence robots and systems, which varied depending on the characteristics of each conference [5].

Other systems, including ThirdEye which presents both first- and third-person perspectives simultaneously to the operator of the telepresence robot [30], and the telepresence robot system to shop with a remote user [31], have been proposed. Additionally, robots with higher degrees of freedom, including MeBot and iRIS, which can change facial orientations and move their arms just like human beings [6], [32]–[35], have also been used in research.

However, these studies have yet to focus primarily on human-robot coexistence. In particular, nonverbal information that was claimed to be necessary for cooperative movement in flat fixed display environments, AR, and VR-based environments is not always conveyed because only what the robot can reproduce can be communicated to the other person.

Therefore, in this study, we use VR technology to give images of people in remote locations to the telepresence robot. By showing the remote worker the image of a person as they are, nonverbal information is communicated.

2.4 Passing by a Telepresence Robot

In general, robots need to ensure the safety of humans [36]. For example, studies have investigated collision avoidance between humans and autonomous robots by using sensors (e.g., [37]–[39]), predicted emotions from a person's face to take avoiding action [40], and designed or developed interfaces for human-operated robots [41]–[43]. However, little is known about situations in which a human passes by a human-operated robot.

Inoue and Yuan have proposed a VR system for smooth physical collaboration between a human and a robot and evaluated the system in a situation in which a human passes a surrogate robot-that is, a robot operated by another worker. The system provided video of the remote worker to a telepresence robot, and the local worker could see the remote worker's body through an HMD. Using an HMD removes restrictions of body movement, and showing the appearance of the remote worker to the local worker prevents them from missing nonverbal information when compared to using avatars [14]. However, their experiment was unrealistic: a worker only walked 1.0 m to pass by the robot. Furthermore, the evaluation was based on a small number of participants.

Sasaki et al. evaluated the VR system through a more realistic experiment with more participants. They reported the results of a questionnaire survey and interviews with participants, which revealed that a worker passing by a robot looked at the human body to predict where the robot was moving to [1]. In this paper, we show results of objective and subjective data, which support the previous claim that a worker looks at the upper body when passing a robot for smooth movement.

3 ENVIRONMENT WHERE A WORKER AND A TELEPRESENCE ROBOT PASS BY EACH OTHER

In this study, we focus on a situation in which a local worker and a telepresence robot operated by a remote worker pass by each other in the same space to investigate how to support the coexistence of humans and robots. In this section, we explain the situation in which a worker and a telepresence robot pass by each other.

3.1 Overview of the Environment

The environment in which a worker and a robot pass by each other is shown in Fig. 3. Two pairs of a worker and a mobile robot equipped with a 360-degree camera (as shown in Fig. 2) were arranged in two separate workspaces. The robots moved according to the position of the remote workers: for example, when worker A moved forward, robot A also moved forward. Each robot moved, synchronizing its respective worker's position, as shown in Fig. 4.

Workers wore HMDs that showed videos captured from the camera of their respective robots; therefore, worker A could see the video of worker B in room Q, acquired by robot A's camera through the HMD, and worker B could see worker A in room P acquired by robot B's camera through the HMD. In this way, an environment was set up in which the local worker and the telepresence robot operated by the remote worker passed by each other. Figure 5 shows an image of the worker in the remote location as presented through the HMD.

Originally, to analyze the system's effectiveness, it was not necessary to prepare two sets of workspaces as shown in Fig. 3, but only one workspace was needed, with one worker, one robot, and one remote worker operating the robot. However, by arranging the worker and robot as shown in Fig. 3, two situations occurred for the two workers, in which the remote worker is operating the robot in front of each worker, and two samples can be obtained at the same time per experiment trial.

3.2 Implementation of the Environment

The environment was constructed as shown in Fig. 3. To eliminate differences between facilities, we set up two rooms

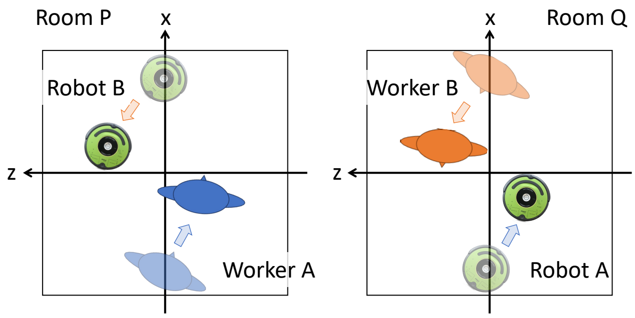


Figure 4: Synchronization between the workers and the robots

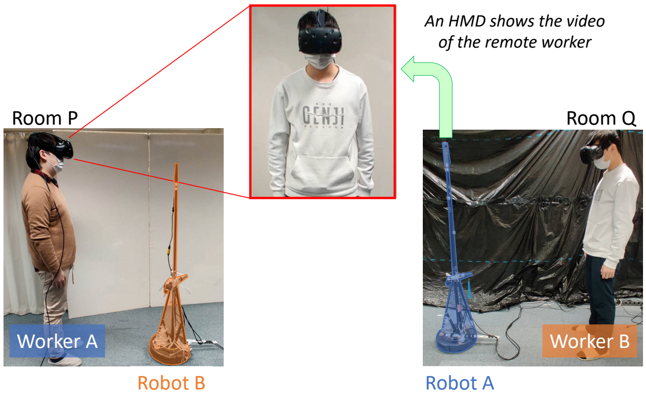


Figure 5: The view in the HMD

of the same size and shape, and the robots and sensors in both rooms were the same. Two tracking cameras were set up on the diagonal of each workspace to establish the positional standard.

The workers and the robots' acquired position and orientation information was sent to a PC, which sent commands to each robot to move, synchronizing the position of the remote workers.

We used VIVE™ products from HTC Corporation[1] in this experiment. The VIVE™ products can obtain position and orientation information by receiving optical lasers from the tracking camera via sensors embedded in the HMD and the controller.

The mobile robot shown in Fig. 2 was the iRobot® Create 2 from iRobot Corporation[2]. iRobot® Create 2 is a mobile robot that can rotate in any direction and move forward and backward. The two wheels on the robot can be commanded separately, and by making one wheel move slower than the other wheel, the robot can rotate while moving forward.

A height-adjustable tripod for the camera was placed on top of the mobile robot. A VIVE™ controller was attached to the tripod. A RICHO[3] 360-degree camera, the Theta V™, was attached to the tip of the tripod to capture first-person images to be presented to the remote worker. The 360-degree camera eliminated the need to move the camera with the worker's orientation, allowing the HMD to adjust its field of view to the worker's small movements.

The images captured by the robot were transmitted to the HMD using WebRTC technology. This entailed a delay from when the camera acquired the image to when it was displayed on the HMD. This delay was an average of 0.43 seconds (SD: 0.10 seconds) based on the preliminary measurements (N=16).

3.3 Synchronization of Worker and Robot Position

In this system, the same systems to match the positions of the worker and the robot were prepared for the two workspaces. Each worker's position was represented by x and z coordinates in Unity™[4]. After obtaining the workers' position information, the PC sent the position coordinates of worker A in the coordinate system of room P to robot A in room Q, and it sent the position coordinates of worker B in the coordinate system of room Q to robot B in room P. Each robot moved based on the location information of the remote worker, as shown in Fig. 4.

4 EXPERIMENT

This study focuses on environments where humans and robots physically coexist and aims to support coexistence. In particular, this study assumes an environment where humans and robots work in the same office. Among the problems that can occur when humans and robots coexist in such an environment, this study focuses on situations where a human and a robot pass by each other in a narrow space, such as a corridor or passageway. In order to support the situation, this study examined the effectiveness of a telepresence robot system developed for real-world cooperative work support [14].

4.1 Passing-by Task

In order to verify how adding a worker's video to each robot would support the situation in which a worker and a robot pass by each other, we experimented with a passing-by task (approved by the Ethics Review Committee of the Faculty of Library, Information and Media Science, University of Tsukuba (Notification number: 20-15)).

In the passing-by task, workers A and B at two remote locations passed by robots B and A, respectively (as shown in Fig. 3), and a previous study [14]. However, in the previous study [14], the initial distance between the worker and the robot was only 0.5 m, and the travel distance was 1.0 m, which was too short for a passing-by experiment. Therefore, in this study, we set the travel distance between the worker and the robot to 3.0 m.

The workspaces were set up in an area of our laboratory. The workspace size was approximately 2.0 m wide by 4.0 m long. Two lines were drawn 3.0 m apart in each workspace;

[1]VIVE - VR Headsets, Games, and Metaverse Life, `https://www.vive.com/` (Visited on Dec 10, 2022)

[2]Coding Robots, Learning Library & STEM Outreach | iRobot Education, `https://edu.irobot.com/` (Visited on Dec 10, 2022)

[3]Ricoh Global | EMPOWERING DIGITAL WORKPLACES, `https://www.ricoh.com/` (Visited on Dec 10, 2022)

[4]Unity Real-Time Development Platform | 3D, 2D VR & AR Engine, `https://unity.com/` (Visited on Dec 10, 2022)

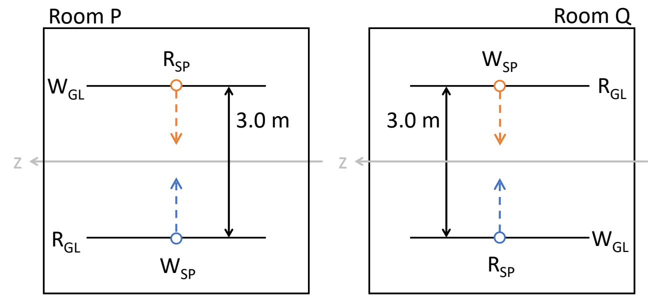


Figure 6: Passing-by task

one indicated the worker's starting point (W_{SP}), and the other indicated the robot's starting point (R_{SP}). We ensured that each line's z coordinates were the same (see Fig. 6). When the experimenter signaled the start, the workers at each starting point started walking toward their goal lines. The robots also moved forward in each workspace because each robot moved in synchronization with a worker's position. The workers and the robots moved forward, slightly shifting directions to avoid collisions. We defined task completion as the time that both workers' bodies crossed their respective worker goal lines (W_{GL}) and both robots' bodies crossed their respective robot goal lines (R_{GL}) in both workspaces. We observed the workers' movements from when they started walking to task completion.

4.2 Experiment Conditions

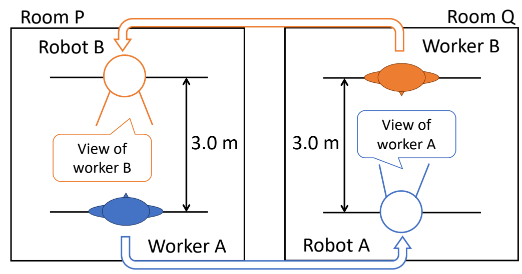
In face-to-face environments, people are known to avoid collisions and coordinate their movements by seeing the other person's actions. Thus the appearance of the person is essential for predicting the other person's movements [9]. For robots and humans to coexist in the same space, methods have also been studied in which the robot captures the human's image with a camera and acts to avoid collisions [39]. However, as we focus on in this study, in an environment where a human and a surrogate robot operated by a human coexist, how a person moves when the other's image, i.e., the robot's operator, has not been studied before, excepting Inoue and Yuan's study [14]. So, we need to reveal the effect of showing the whole body to a person in an environment with a telepresence robot.

In this study, we conducted an experiment to investigate the effect of projecting the operator's entire body image onto a robot through an HMD when the human and the telepresence surrogate robot pass by each other in an environment where they coexist.

We compared two conditions in the experiment based on what workers could see through the HMDs during the passing-by task. One is the **human condition**; in this condition, as described thus far, workers were wearing HMDs at both remote locations, and each HMD displayed the video from the camera mounted on the robot (as shown in Fig. 7).

The other is the **robot condition**; in this condition, worker A's HMD displayed the see-through video in front of them (as shown in Fig. 8). Unlike the human condition, worker A had to move while looking at robot B, although worker B moved

(a)



(b)

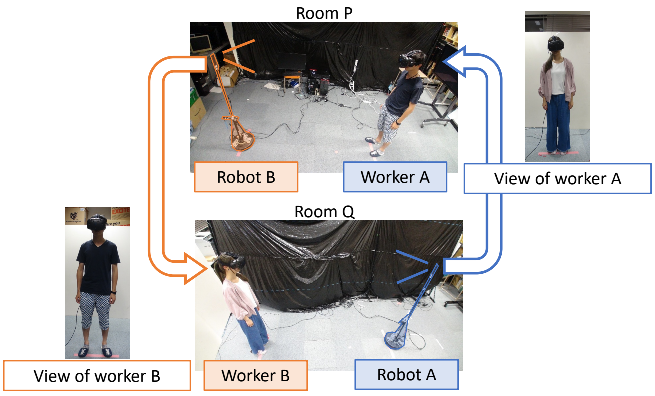


Figure 7: The human condition: (a) an overview; (b) the view of each worker.

while looking at worker A to operate robot B during the task. So, situations were not symmetrical between the participants in a pair. Therefore, only worker A, who performed the task while looking at the robot, was subject to the analysis, and worker B was not. Robot A, which passed by worker B and was operated by worker A, was not used in the robot condition.

In both conditions, workers were not allowed to talk to each other to stimulate visual information transfer between workers' movements. Therefore, during the experiment, workers moved based only on visual information. In comparing the two conditions, we analyzed the effect of displaying the remote worker's full-body image to the local worker during the passing-by task.

4.3 Participants

Six pairs of 12 graduate students (all male) participated in the experiment. Of the 12, five had never used an HMD before this experiment. Only one of the six pairs had never met each other before.

4.4 Procedure

The experiment consisted of a within-participant design; Table 1 shows the experiment's procedure. First, each pair was assigned to the human or robot condition, which were counterbalanced. In each condition, participants were assigned to be workers A or B and then moved to the remote workspaces rooms P or Q. After the experimenter explained the system, the participants practiced the passing-by

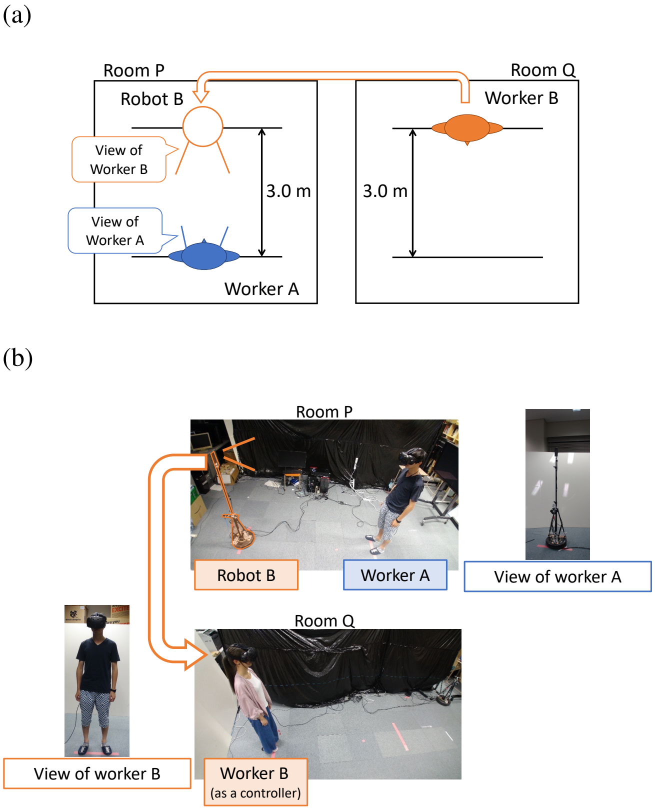


Figure 8: The robot condition: (a) an overview; (b) the view of each worker.

task twice to familiarize themselves with the task and use of the system. The experimenter delineated the placement of the goal line and how far the participants could move in the workspace, explaining until both participants understood the task. The experimenter also asked participants to walk without bumping into the robot and did not mention how fast or long they should walk.

After the practice rounds, the participants performed the task once. After the task, the sequence was repeated under another condition. The robot condition differed from the human condition in that worker A passed by robot B, but worker B walked alone-that is, the robot condition was asymetrical. Therefore, after completing the task, workers A and B switched roles and performed the task (including practice rounds) again. As a result, the task was performed a total of six times (one time for each pair) for all pairs in the human condition and 12 times (two times for each pair) in the robot condition (excluding the practices).

One experimenter was assigned to room P, and another was assigned to room Q, to prevent workers from colliding with the robot and from catching their feet on the cords connecting the robot and the PC during the experiment.

A questionnaire survey was conducted after the task in each condition. According to previous studies investigating remote cooperative work through robot assistance, whether workers are satisfied with remote cooperative work involving robots depends on the presence of the partners and the feeling that they are collaborating with the partners [44], [45]. Therefore, in this study, ten questions involving subjective evaluations were posed, asking not only about the robot's usability but also about the presence of the partner and the sense of being in the same room with the robot. Participants responded to the evaluation items using a 7-point scale ranging from "strongly disagree" to "strongly agree." The questionnaire items are shown in Table 2.

In both conditions, we also conducted open-ended interviews after the task was completed. Participants were asked to respond freely to questions about their impressions of the task and the system.

4.5 Data Collection

The coordinates of the workers and the robots were acquired using HMDs worn by workers and controllers mounted on mobile robots. In the human condition, coordinates were acquired for six workers and six robots in six pairs and two locations; in the robot condition, coordinates were acquired for 12 workers and 12 robots in six pairs and one location (Room P). The frequency of coordinate capturing was five times per second (every 0.2 seconds), the same as the frequency at which coordinates were exchanged between the HMD and the PC.

To examine how the workers passed by each other in response to the visual information presented, we obtained the video that the HMD showed the workers during the task. HMD videos were not gathered during practice rounds, so there were 12 videos for six pairs in two rooms in the human condition and 12 videos for six pairs in one room (Room P) performing the task two times in the robot condition. Thus, a total of 24 videos were obtained across both conditions.

The questionnaire survey results and the interview transcripts were also used for subjective evaluation.

5 RESULT

In the experiments conducted in this study, there was the possibility of collisions between the robot and the worker. However, no cases of collisions occurred between the robot and the worker. There were two cases in which a worker's leg was touched lightly by the robot in the human condition, but since none of these cases resulted in worker injury, the experiment was not interrupted.

In this chapter, we present the results of our analysis of the quantitative data (the users' movements and the direction of the users' field of view) and of the qualitative data (the questionnaire evaluation) [1].

5.1 The Workers' Movement

Table 3 shows the mean values of the workers' walking time, workers' walking distance, workers' walking speed, robots' moving distance, robots' moving speed during the task, and time from commencement of walking to passing the robot (N=12). A one-way MANOVA (two levels of condition: human condition and robot condition) with the dependent variables of worker walking time, worker walking distance, worker walking speed, robot moving distance, robot moving speed, and time from commencement of walking to passing the robot was conducted to test differences in the

Table 1: The experiment procedure

Procedure	Human condition	Robot condition
1.	Workers A and B stood on their respective start points (see 7(a)).	Workers A and B stood on their respective start points (see Fig. 8(a)).
2.	Workers were informed of the system and practiced the task twice.	Workers were informed of the system and practiced the task twice.
3.	Workers performed the task once.	Workers performed the task once.
4.	Workers filled out the questionnaire.	Worker A filled out the questionnaire.
5.	-	The role of workers A and B were switched, and they stood at their respective start points.
6.	-	Workers were informed of the system and practiced the task twice.
7.	-	Workers performed the task once.
8.	-	Worker A filled out the questionnaire.
9.	Workers were interviewed after both conditions had been conducted to completion.	

Table 2: Questionnaire items

	Items
Q1	After using the system, I felt uncomfortable in a way similar to motion sickness.
Q2	The video was clear and pretty.
Q3	I felt like the remote person was in the same room with me.
Q4	I could clearly perceive the motion of my partner.
Q5	I felt like I passed the other person in the same room.
Q6	The field of view changed naturally in line with my movements.
Q7	I perceived the movements and directions of my partner.
Q8	I could predict where my partner was moving to.
Q9	My partner's movement looked realistic.
Q10	The experience in the virtual environment was consistent with experiences in real life.

means between the conditions, but no significant differences were found ($F(1, 6) = 0.37,\ p = .89$).

5.2 Workers' Fields of View During the Task

Figure 9 shows an example of the video displayed on the HMD during the experiment, which the worker was viewing. This image presents the worker's field of view (FOV), which is categorized into three patterns based on where the center of the FOV (red dot in Fig. 9) was facing: the upper body, the lower body, or the floor. The definitions of the direction of the worker's FOV are shown in Table 4. For example, when the center of the worker's FOV was above the waist in the human condition, the worker was considered to be looking at the upper body; in the robot condition, when the center of the worker's FOV was facing the drive part, the worker was considered to be looking at the lower body. The measurements were taken from the start of walking to the point of passing each other. The worker and the robot faced either the upper body, the lower body, or the floor, and they never faced another direction until they passed each other. The average time spent looking at each part is shown in Table 4, and the average ratio for all pairs is shown in Fig. 10.

Table 3: Averages of movements of workers and robots

	Human condition	Robot condition
Walking time [s]	12.28 (SD: 3.04)	11.67 (SD: 3.61)
Walking distance of the worker [m]	3.64 (SD: 0.29)	3.67 (SD: 0.38)
Walking speed of the worker [cm/s]	31.2 (SD: 6.96)	33.7 (SD: 8.19)
Moving distance of the robot [m]	2.94 (SD: 0.58)	2.73 (SD: 0.78)
Moving speed of the robot [cm/s]	24.3 (SD: 2.72)	24.4 (SD: 3.50)
Time from starting walking to passing by the robot [s]	7.27 (SD: 1.48)	7.30 (SD: 2.91)

The one-way MANOVA (two levels of condition: human condition and robot condition) conducted with the dependent variable of the ratio of time spent looking at each viewing body part showed significant differences ($F(1, 22) = 11.30,\ p = .0002$). We also conducted corresponding Bonferroni t-tests, which revealed significant differences in the following six items (as shown in Fig. 10):

- Time spent looking at the upper body in the human condition vs. time spent looking at the lower body in the human condition ($t(11) = 7.16,\ p = .0003$)
- Time spent looking at the upper body in the human condition vs. time spent looking at the floor in the human condition ($t(11) = 11.38,\ p < .01$)
- Time spent looking at the upper body in the human condition vs. time spent looking at the upper body in the robot condition ($t(11) = 5.73,\ p = .002$)
- Time spent looking at the upper body in the human condition vs. time spent looking at the lower body in the robot condition ($t(11) = 4.37,\ p = .02$)
- Time spent looking at the upper body in the human condition vs. time spent looking at the floor in the robot

Table 4: Time viewing body parts

	Human condition		Robot condition	
Definition	Direction of the center of the FoV	Mean/SD [s]	Direction of the center of the FoV	Mean/SD [s]
Upper body	Above the waist	6.88 / 2.01	Above the center of the robot (including the prop and 360-degree camera)	1.40 / 2.24
Lower body	Below the waist and above the floor	0.38 / 1.27	Below the center and above the floor of the robot (including the prop)	3.25 / 2.40
Floor	Floor	0.00 / 0.00	Floor	2.65 / 2.94

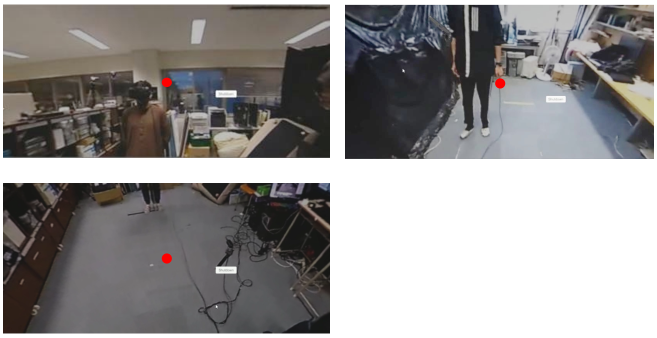


Figure 9: Screenshots displayed to workers. The upper left shows the worker looking at the other worker's upper body; the upper right shows the worker looking at the other worker's lower body; the lower left shows the worker looking at the floor.

condition ($t(11) = 5.28,\ p = .004$)

- Time spent looking at the floor in the human condition vs. time spent looking at the lower body in the robot condition ($t(11) = -4.50,\ p = .01$)

In particular, the first five items indicated that more time is spent looking at the upper body in the human condition than in either of the two conditions, implying that the worker is looking at the upper body longer because they can see the human image.

5.3 Questionnaire

Ratings were treated as a score from 1 to 7, and a Wilcoxon signed-rank test was conducted for each question item; the results are shown in Table 5. In particular, the scores of four questions were significantly higher in the human condition than in the robot condition: Q4 "I could clearly perceive the motion of my partner" ($p < .01$); Q7 "I perceived the movements and directions of my partner" ($p < .01$); Q8 "I could predict where my partner was moving to" ($p < .01$); and Q9 "My partner's movement looked realistic" ($p < .05$).

6 DISCUSSION

In the human and the robot conditions, a significant difference was found in the direction of the worker's face during the task: workers in the human condition spent relatively more time looking at the upper body of the remote worker, while workers in the robot condition spent more time looking at the lower body and the floor. First, we discuss the results of the questionnaire survey and the interviews.

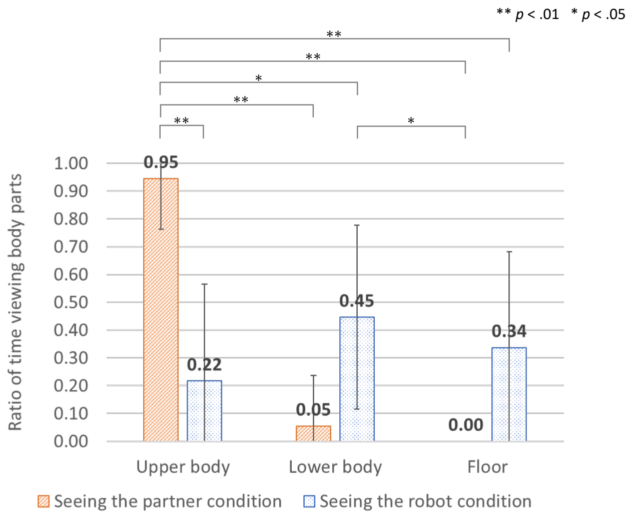


Figure 10: Time viewing body parts. Asterisk(s) show significant differences, and error bars show standard deviations.

In the questionnaire survey, as mentioned in section 5.3, there were significant differences in the following items: Q4 "I could clearly perceive the motion of the partner", Q7 "I perceived the movements and directions of my partner", and Q8 "I could predict where my partner was moving to." These differences are considered to have occurred because the human whole-body image was presented exactly as the remote robot's camera captured it in the human condition.

In addition, in the interviews, participants in the human condition provided such comments as "I could predict the robot's path when looking at a human (P3)" and "I could clearly see its [the robot's] direction of movement when looking at a human (P5)." In contrast, participants in the robot condition provided these comments: "I could not see the robot's direction of movement (P2, P7)" and "I was confused about the robot's direction of movement (P8)." Based on the comments' differences, the local workers seemed to predict the robot's movement in the human condition, in which the human image was seen in the HMD. In contrast, the robot's direction of movement was unpredictable in the robot condition, in which the robot's image was seen in the HMD.

The result of the interviews suggests that the worker in the

Table 5: The study's questionnaire; p-values are the results of Wilcoxon signed-rank test.

($^{**}\ p < .01,\ ^{*}\ p < .05,\ ^{\dagger}\ p < .10$)

	Items	Mean / Median / Mode in the human condition	Mean / Median / Mode in the robot condition	p-value
Q1	After using the system, I felt uncomfortable in a way similar to motion sickness.	4.0 / 5 / 2, 5, 6	4.3 / 5 / 5	.79
Q2	The video was clear and pretty.	5.1 / 5 / 5	4.3 / 4.5 / 3, 5	$.08^{\dagger}$
Q3	I felt like the remote person was in the same room with me.	6.3 / 6.5 / 7	5.5 / 6 / 6	.13
Q4	I could clearly perceive the motion of my partner.	6.3 / 6 / 6	5.3 / 5 / 5	$.008^{**}$
Q5	I felt like I passed the other person in the same room.	5.3 / 5.5 / 5, 6, 7	5.6 / 5.5 / 5	.69
Q6	The field of view changed naturally in line with my movements.	5.5 / 5.5 / 5	5.4 / 5.5 / 5	1.00
Q7	I perceived the movements and directions of my partner.	6.1 / 6 / 6	4.6 / 4.5 / 4	$.003^{**}$
Q8	I could predict where my partner was moving to.	6.2 / 6 / 6	4.2 / 4 / 4	$.001^{**}$
Q9	My partner's movement looked realistic.	5.8 / 6 / 6	4.2 / 4.5 / 5	$.02^{*}$
Q10	The experience in the virtual environment was consistent with experiences in real life.	5.6 / 5.5 / 5	4.8 / 5 / 4, 5	.16

human condition predicted the robot's next move by looking at the upper body but that the worker in the robot condition could not predict the robot's movement by looking at the upper body. This may be why the worker looked at the lower body, which is the robot's moving part, longer in the robot condition to try to predict where the robot would go. This indicates that presenting an image of the worker's upper body is important to predict the remote worker's action when cooperative work involving movement is performed across remote locations.

Previous studies have also proven the importance of having the human image in collaborative work. A study of a collaborative bicycle repair task in face-to-face and remote contexts found that the visual cues in the human image could provide the grounding for their conversation and the task [10]. Another study involving a remote collaborative task of assembling a toy was conducted in a specially designed space where the local worker and a table were surrounded by eight fixed displays showing a remote worker; it found that the local worker could predict the remote worker's movement by seeing the upper body of their remote worker and could thus prepare for the following action [11]. Unlike these studies of predefined collaborative work in a fixed place, our study demonstrates the usefulness of showing the human video when passing a robot, as this helps predict the robot's moving direction.

So far, we have mentioned that by projecting human images onto a robot, a person can predict the robot's direction of travel. We will further discuss other effects of human images on robots.

Considering the worker's FOV, a robot with an image of a person can not only predict the actions of the remote worker but also prevent the worker from being unable to see its surroundings by looking down. This study considered a situation in which workers and surrogate robots worked together in the same workspace. Walking while looking down can lead to a collision with a robot, another person, or an object, not only in a narrow spaces but also in other spaces in an office. Presenting a human video can prevent collision with others during remote cooperative work.

It is important to see the other person passing by when passing each other. However, the robot with no person's image lost various sources of human information, such as body movements and gaze, which may cause a certain sense of anxiety that the human may collide with the robot. Conversely, giving a human image to a robot may help remove this insecurity and give a sense of safety, such as "I can avoid collision with the robot."

It should be noted that this study has limitations.

In this experiment, for safety reasons, robots with a low center of gravity and concentrated major parts (including the drive unit on the bottom surface) were used, which is the same model as the cleaning robots that are currently available in relatively large numbers. Since this mobile robot has a geometrically symmetrical shape, it is difficult to determine the direction in which it is moving, especially when looking at the figure of the robot, as in the robot condition. If this were a humanoid robot in which the direction of movement was more perceptible than a symmetrical-shaped robot, the experimental results of this study might be different. In any case, the findings of this study may be helpful in building a telepresence robot system using current major mobile robots, like the one used in this study.

In the current experiment, the workers' walking speed was approximately 30 cm/s, which is slower than the speed at which humans generally walk. Various factors should be considered:

- The workspace was small.
- The worker's foot could have caught on the cord connecting the robot and the PC.
- Two practice rounds were insufficient for confident walking with the HMD.

We hope to study and analyze the results of higher walking speeds in the future.

Finally, we discuss applying our findings to the real world. This study assumes an in-company environment where people and robots come and go freely. Therefore, besides a corridor, people and robots move in various places and situations, and it is hard to establish rules such as "any persons must keep to the left, and any robots must keep to the right." In previous studies in which robots moved in office environments or other environments with humans (e.g., [46] and [47]), the experiments were conducted in environments that allowed movement within the same space without restrictions on movement for both the persons and the robot. As in these studies, in studying ways for humans and robots to coexist, it is also necessary to consider methods that do not depend on rules. Thus, our findings on the effectiveness of showing the person's whole body may be helpful in situations where humans and robots coexist.

However, preparing the equipment is more burdensome than in a conventional environment. In our environment, the robots need to acquire the position of a remote person and show the remote person to each other. We used sensors to realize the former and HMDs to realize the latter. In the future, it will be possible to reduce the burden of system construction and the user by using a camera attached to the HMD to acquire location information through image recognition or by using AR glasses, which are lighter and less burdensome for the user. Additionally, to use our findings, a space for the robot operator to walk around is necessary. This research scenario is feasible if, for example, a satellite office or coworking space is used as a remote office. In addition, these devices that can freely walk in place in 360 degrees (e.g., the systems of [48], [49], or the Virtuix™ Omni and Cyberith Virtualizer used in [50]) can be used to realize this environment in a smaller physical space.

7 CONCLUSION

The purpose of this study is to support human-robot coexistence. Assuming an office where humans and telepresence surrogate robots work in the same place, we focused on a situation in which a human and a robot pass by each other. This study uses a VR system that presents a remote user's whole-body image to the robot [14], which was initially proposed to support real-world cooperative work between a human and a robot. We verified how the system could support the passing-by of a human and a robot. We conducted an experiment and analyzed subjective and objective data from the passing-by task. The results show that workers using the system looked at the human upper body to predict the robot's movement. This indicates that in an environment where humans and robots coexist, projecting a full-body image of a person onto a robot is helpful for smoothly moving in the same physical space.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Yasuhito Noguchi for his cooperation in conducting the experiment.

REFERENCES

[1] Sasaki, K., Yuan, Z. and Inoue, T.: An Experiment of an Office Worker Passing by the Surrogate Robot of a Remote Worker Using VR Video, *Proceedings of the 16th International Workshop on Informatics (IWIN 2022)*, pp. 133–137 (2022).

[2] Shen, S., Tennent, H., Claure, H. and Jung, M.: My Telepresence, My Culture? An Intercultural Investigation of Telepresence Robot Operators' Interpersonal Distance Behaviors, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, Association for Computing Machinery, p. 1–11 (online), https://doi.org/10.1145/3173574.3173625 (2018).

[3] Myodo, E., Xu, J., Tasaka, K., Yanagihara, H. and Sakazawa, S.: Issues and solutions to informal communication in working from home using a telepresence robot, *ITE Transactions on Media Technology and Applications*, Vol. 6, No. 1, pp. 30–45 (2018).

[4] Ahumada-Newhart, V. and Olson, J. S.: Going to School on a Robot: Robot and User Interface Design Features That Matter, *ACM Trans. Comput.-Hum. Interact.*, Vol. 26, No. 4 (online), https://doi.org/10.1145/3325210 (2019).

[5] Neustaedter, C., Singhal, S., Pan, R., Heshmat, Y., Forghani, A. and Tang, J.: From Being There to Watching: Shared and Dedicated Telepresence Robot Usage at Academic Conferences, *ACM Trans. Comput.-Hum. Interact.*, Vol. 25, No. 6 (online), https://doi.org/10.1145/3243213 (2018).

[6] Nagendran, A., Steed, A., Kelly, B. and Pan, Y.: Symmetric telepresence using robotic humanoid surrogates, *Computer Animation and Virtual Worlds*, Vol. 26, No. 3-4, pp. 271–280 (2015).

[7] Guizzo, E.: When my avatar went to work, *IEEE Spectrum*, Vol. 47, No. 9, pp. 26–50 (2010).

[8] Tachi, S., Kawakami, N., Nii, H., Watanabe, K. and Minamizawa, K.: Telesarphone: Mutual telexistence master-slave communication system based on retroreflective projection technology, *SICE Journal of Control, Measurement, and System Integration*, Vol. 1, No. 5, pp. 335–344 (2008).

[9] Kendon, A.: *Conducting interaction: Patterns of behavior in focused encounters*, Vol. 7, CUP Archive (1990).

[10] Fussell, S. R., Kraut, R. E. and Siegel, J.: Coordination of Communication: Effects of Shared Visual Context on Collaborative Work, *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*, CSCW '00, New York, NY, USA, Asso-

ciation for Computing Machinery, pp. 21–30 (online), https://doi.org/10.1145/358916.358947 (2000).

[11] Yamashita, N., Kuzuoka, H., Hirata, K., Aoyagi, S., Shirai, Y., Kaji, K. and Harada, Y.: Effects of Showing User's Upper Body in Video-mediated Collaboration, *IPSJ Journal*, Vol. 51, No. 4, pp. 1152–1162 (online), https://ci.nii.ac.jp/naid/110007970719/en/ (2010).

[12] Burgoon, J. K., Birk, T. and Pfau, M.: Nonverbal behaviors, persuasion, and credibility, *Human communication research*, Vol. 17, No. 1, pp. 140–169 (1990).

[13] Sundaram, D. S. and Webster, C.: The role of nonverbal communication in service encounters, *Journal of Services Marketing* (2000).

[14] Inoue, T. and Yuan, Z.: Remote Video Figure Achieves Smooth Cooperative Movement in a Bidirectional Telepresence Robot Environment, *Collaboration and Technology* (Rodrigues, A., Fonseca, B. and Preguiça, N., eds.), Cham, Springer International Publishing, pp. 91–104 (2018).

[15] Jones, A., Lang, M., Fyffe, G., Yu, X., Busch, J., McDowall, I., Bolas, M. and Debevec, P.: Achieving Eye Contact in a One-to-Many 3D Video Teleconferencing System, *ACM Trans. Graph.*, Vol. 28, No. 3 (online), https://doi.org/10.1145/1531326.1531370 (2009).

[16] Tang, J. C. and Minneman, S.: VideoWhiteboard: Video Shadows to Support Remote Collaboration, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, New York, NY, USA, Association for Computing Machinery, pp. 315–322 (online), https://doi.org/10.1145/108844.108932 (1991).

[17] Handberg, L., Gullströ, C., Kort, J. and Nyström, J.: Spatial and Social Connectedness in Web-Based Work Collaboration, *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*, CSCW '16 Companion, New York, NY, USA, Association for Computing Machinery, pp. 45–48 (online), https://doi.org/10.1145/2818052.2874321 (2016).

[18] Inoue, T., Nawahdah, M. and Noguchi, Y.: User's Communication Behavior in a Pseudo Same-room Videoconferencing System BHS, *International Journal of Informatics Society (IJIS)*, Vol. 6, pp. 39–47 (2014).

[19] Hirata, K., Harada, Y., Takada, T., Aoyagi, S., Shirai, Y., Yamashita, N., Kaji, K., Yamato, J. and Nakazawa, K.: t-Room: Next Generation Video Communication System, *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, pp. 1–4 (online), https://doi.org/10.1109/GLOCOM.2008.ECP.1058 (2008).

[20] Pejsa, T., Kantor, J., Benko, H., Ofek, E. and Wilson, A.: Room2Room: Enabling Life-Size Telepresence in a Projected Augmented Reality Environment, *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, New York, NY, USA, Association for Computing Machinery, pp. 1716–1725 (online), https://doi.org/10.1145/2818048.2819965 (2016).

[21] Maimone, A., Yang, X., Dierk, N., State, A., Dou, M. and Fuchs, H.: General-purpose telepresence with head-worn optical see-through displays and projector-based lighting, *2013 IEEE Virtual Reality (VR)*, pp. 23–26 (online), https://doi.org/10.1109/VR.2013.6549352 (2013).

[22] Fuchs, H., State, A. and Bazin, J.: Immersive 3D Telepresence, *Computer*, Vol. 47, No. 7, pp. 46–52 (online), https://doi.org/10.1109/MC.2014.185 (2014).

[23] You, B.-J., Kwon, J. R., Nam, S.-H., Lee, J.-J., Lee, K.-K. and Yeom, K.: Coexistent Space: Toward Seamless Integration of Real, Virtual, and Remote Worlds for 4D+ Interpersonal Interaction and Collaboration, *SIGGRAPH Asia 2014 Autonomous Virtual Humans and Social Robot for Telepresence*, SA '14, New York, NY, USA, Association for Computing Machinery, (online), https://doi.org/10.1145/2668956.2668957 (2014).

[24] Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., Kim, D., Davidson, P. L., Khamis, S., Dou, M. and et al.: Holoportation: Virtual 3D Teleportation in Real-Time, *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, UIST '16, New York, NY, USA, Association for Computing Machinery, pp. 741–754 (online), https://doi.org/10.1145/2984511.2984517 (2016).

[25] Komiyama, R., Miyaki, T. and Rekimoto, J.: JackIn Space: Designing a Seamless Transition between First and Third Person View for Effective Telepresence Collaborations, *Proceedings of the 8th Augmented Human International Conference*, AH '17, New York, NY, USA, Association for Computing Machinery, (online), https://doi.org/10.1145/3041164.3041183 (2017).

[26] Lee, G. A., Teo, T., Kim, S. and Billinghurst, M.: Mixed Reality Collaboration through Sharing a Live Panorama, *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*, SA '17, New York, NY, USA, Association for Computing Machinery, (online), https://doi.org/10.1145/3132787.3139203 (2017).

[27] Piumsomboon, T., Day, A., Ens, B., Lee, Y., Lee, G. and Billinghurst, M.: Exploring Enhancements for Remote Mixed Reality Collaboration, *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*, SA '17, New York, NY, USA, Association for Computing Machinery, (online), https://doi.org/10.1145/3132787.3139200 (2017).

[28] Piumsomboon, T., Lee, G. A., Hart, J. D., Ens, B., Lindeman, R. W., Thomas, B. H. and Billinghurst, M.: Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, Association for Computing Machinery, (online), https://doi.org/10.1145/3173574.3173620 (2018).

[29] Teo, T., F. Hayati, A., A. Lee, G., Billinghurst, M. and Adcock, M.: A Technique for Mixed Reality Remote Collaboration Using 360 Panoramas in 3D Reconstructed Scenes, *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19, New York, NY, USA, Association for Computing Machinery, (on-

line), https://doi.org/10.1145/3359996.3364238 (2019).

[30] Tarun, A. P., Baig, N. M., Chang, J. S.-K., Tanvir, R., Shihipar, S. and Mazalek, A.: Third Eye: Exploring the Affordances of Third-Person View in Telepresence Robots, *Social Robotics* (Salichs, M. A., Ge, S. S., Barakova, E. I., Cabibihan, J.-J., Wagner, A. R., Castro-González, Á. and He, H., eds.), Cham, Springer International Publishing, pp. 707–716 (2019).

[31] Yang, L., Jones, B., Neustaedter, C. and Singhal, S.: Shopping Over Distance through a Telepresence Robot, *Proc. ACM Hum.-Comput. Interact.*, Vol. 2, No. CSCW (online), https://doi.org/10.1145/3274460 (2018).

[32] Adalgeirsson, S. O. and Breazeal, C.: MeBot: A robotic platform for socially embodied telepresence, *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 15–22 (online), https://doi.org/10.1109/HRI.2010.5453272 (2010).

[33] Kawanobe, H., Aosaki, Y., Kuzuoka, H. and Suzuki, Y.: iRIS: A remote surrogate for mutual reference, *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 403–403 (online), https://doi.org/10.1109/HRI.2013.6483618 (2013).

[34] Van Erp, J. B., Sallaberry, C., Brekelmans, C., Dresscher, D., Ter Haar, F., Englebienne, G., Van Bruggen, J., De Greeff, J., Pereira, L. F. S., Toet, A., Hoeba, N., Lieftink, R., Falcone, S. and Brug, T.: What Comes After Telepresence? Embodiment, Social Presence and Transporting One's Functional and Social Self, *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2067–2072 (online), https://doi.org/10.1109/SMC53654.2022.9945544 (2022).

[35] Nakanishi, H., Tanaka, K. and Wada, Y.: Remote Handshaking: Touch Enhances Video-Mediated Social Telepresence, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, New York, NY, USA, Association for Computing Machinery, pp. 2143–2152 (online), https://doi.org/10.1145/2556288.2557169 (2014).

[36] Claes, D. and Tuyls, K.: Multi robot collision avoidance in a shared workspace, *Autonomous Robots*, Vol. 42 (online), https://doi.org/10.1007/s10514-018-9726-5 (2018).

[37] Prassler, E., Scholz, J. and Fiorini, P.: Navigating a Robotic Wheelchair in a Railway Station during Rush Hour, *The International Journal of Robotics Research*, Vol. 18, No. 7, pp. 711–727 (online), https://doi.org/10.1177/02783649922066529 (1999).

[38] Clodic, A., Fleury, S., Alami, R., Chatila, R., Bailly, G., Brethes, L., Cottret, M., Danes, P., Dollat, X., Elisei, F., Ferrane, I., Herrb, M., Infantes, G., Lemaire, C., Lerasle, F., Manhes, J., Marcoul, P., Menezes, P. and Montreuil, V.: Rackham: An Interactive Robot-Guide, *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 502–509 (online), https://doi.org/10.1109/ROMAN.2006.314378 (2006).

[39] Ratsamee, P., Mae, Y., Ohara, K., Takubo, T. and Arai, T.: Human–robot collision avoidance using a modified social force model with body pose and face orientation, *International Journal of Humanoid Robotics*, Vol. 10, No. 01, p. 1350008 (2013).

[40] Bera, A., Randhavane, T. and Manocha, D.: The emotionally intelligent robot: Improving socially-aware human prediction in crowded environments, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, p. 6 (2019).

[41] Gomez, C., Hernandez, A., Crespo, J. and Barber, R.: LEARNING ROBOTICS THROUGH A FRIENDLY GRAPHICAL USER INTERFACE, *ICERI2016 Proceedings*, 9th annual International Conference of Education, Research and Innovation, IATED, pp. 483–492 (online), https://doi.org/10.21125/iceri.2016.1120 (2016).

[42] Ajmera, Y. and Javed, A.: Shared autonomy in web-based human robot interaction, *Proceedings of SAI Intelligent Systems Conference*, Springer, pp. 696–702 (2020).

[43] Birkenkampf, P., Leidner, D. and Borst, C.: A knowledge-driven shared autonomy human-robot interface for tablet computers, *2014 IEEE-RAS International Conference on Humanoid Robots*, pp. 152–159 (online), https://doi.org/10.1109/HUMANOIDS.2014.7041352 (2014).

[44] Tanaka, K., Wada, Y. and Nakanishi, H.: Remote Handshaking: Integration of a Haptic Device and Videoconferencing to Enhance Social Telepresence, Technical Report 4 (2015).

[45] Murakami, Y., Nakanishi, H., Nogami, D. and Ishiguro, H.: Effects of Robot-mounted Cameras' Movement on Telepresence, *IPSJ Journal*, Vol. 51, No. 1, pp. 54–62 (online), `https://ci.nii.ac.jp/naid/110007970615/en/` (2010).

[46] Asoh, H., Motomura, Y., Asano, F., Hara, I., Hayamizu, S., Itou, K., Kurita, T., Matsui, T., Vlassis, N., Bunschoten, R. et al.: Jijo-2: An office robot that communicates and learns, *IEEE Intelligent Systems*, Vol. 16, No. 5, pp. 46–55 (2001).

[47] Severinson-Eklundh, K., Green, A. and Hüttenrauch, H.: Social and collaborative aspects of interaction with a service robot, *Robotics and Autonomous systems*, Vol. 42, No. 3-4, pp. 223–234 (2003).

[48] Templeman, J. N., Denbrook, P. S. and Sibert, L. E.: Virtual locomotion: Walking in place through virtual environments, *Presence*, Vol. 8, No. 6, pp. 598–617 (1999).

[49] Nakamura, J., Matsuda, Y., Amemiya, T., Ikei, Y. and Kitazaki, M.: Virtual Walking With Omnidirectional Movies and Foot Vibrations: Scene-Congruent Vibrations Enhance Walking-Related Sensations and Ground Material Perceptions, *Ieee Access*, Vol. 9, pp. 168107–168120 (2021).

[50] Hooks, K., Ferguson, W., Morillo, P. and Cruz-Neira, C.: Evaluating the user experience of omnidirectional VR walking simulators, *Entertainment Computing*, Vol. 34, p. 100352 (2020).

(Received: December 13, 2022)
(Accepted: August 30, 2023)

Kosuke Sasaki received his M.Sc. in Informatics from the University of Tsukuba in 2016. He is currently studying in the University of Tsukuba, Japan. He is also a Specially Appointed Assistant Professor of the Faculty of Global Management, Chuo University, Japan. His research interests include Computer-Supported Cooperative Work and Educational Technology.

Zijie Yuan received his M.Sc. in Informatics from the University of Tsukuba in 2020. He currently works for Rakuten Mobile, Japan. His research interests include Computer-Supported Cooperative Work and mobile computing application.

Tomoo Inoue is Professor of the Institute of Library, Information and Media Science of University of Tsukuba, Japan. His research interests include Computer-Supported Cooperative Work, Human-Computer Interaction and Educational Technology. He received his Ph.D. in Engineering from Keio University in 1998. He is a recipient of awards including Outstanding Paper Award, Activity Contribution Award from Information Processing Society of Japan (IPSJ), and Best Faculty Member of 2021 from his university. He has served a number of academic committees, including IPSJ Journal Group Editor-in-Chief, IPSJ Transactions on Digital Content Editor-in-Chief, IEICE Technical Committee on Human Communication Science Chair, IPSJ SIG Groupware and Network Services Chair, ACM CSCW Papers Associate Chair, CollabTech Steering Committee, and IEEE TC CSCWD.

Regular Paper

Prototype Development and Evaluation of an Avatar Remote Communication System for ALS Patients Using Eye-tracking Interface

Nagisa Kokubu [*], Takayo Namba [**], and Keiichi Abe [*]

[*] Kanagawa Institute of Technology, Atsugi City, Kanagawa, Japan
[**] Kawasaki City College of Nursing, Kawasaki City, Kanagawa, Japan

abe@he.kanagawa-it.ac.jp

Abstract -ALS patients with severe disease have difficulty in walking, going out, and talking with people. The authors believe that there is a need for a system that enables people with severe ALS to communicate with many people at home while maintaining contact with society. A typical example of a solution to this problem is an alter-ego robot. This robot is placed in a place where ALS patients cannot go, and it can communicate with the other person remotely by gestures. However, the robot needs to be set up at the place where the patient is going to talk with the other person, and the robot needs to be removed from the place after the conversation. In this study, we propose a system that enables ALS patients with severe symptoms who have difficulty in going out to talk with many people through avatars such as computer graphics (CG), which are alter ego characters, easily and remotely by the patients themselves. We also describe in detail the development and evaluation of a prototype of our proposed system.

Keywords: Remote Communication, ALS, online conferencing tool.

1 INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a disease in which the nerves that control voluntary movements (motor neurons) are damaged, resulting in the loss of muscles throughout the body and those necessary for conversation. In severe cases, the patient becomes bedridden and isolated, with no contact with other people, and it is difficult for the patient to move his/her body, although he/she is conscious. A previous study aiming to solve this problem is an avatar robot [1]. This avatar robot is placed in a place where ALS patients cannot go, and it can communicate with a person remotely by gestures. However, it is necessary to set up the avatar robot and remove it after the conversation. Therefore, the system requires help from other people besides the ALS patient.

In this study, we propose a system that enables ALS patients to communicate with many people by themselves in order to keep contact with society and enjoy communication with various people. In this study, we propose a system that enables ALS patients to talk with many people at their own will by using eye control technology, avatar technology, and online conferencing tools [2]. Furthermore, the system allows ALS patients to choose their favorite avatar character during the conversation. We developed and evaluated a prototype of the proposed system. In order to evaluate the effectiveness of the proposed system from various viewpoints, we performed three questionnaires: one is a questionnaire for ALS patients, two is a questionnaire for medical professionals who actually work with ALS patients. and the other is a questionnaire for able-bodied persons. The details are described below.

2 RELARED WORKS

When ALS patients become severely ill, it becomes difficult for them to move their bodies and to speak. Conversation aids are an existing technology to help severely ill ALS patients communicate their intentions. The first representative product of portable conversational aids is a keyboard input type conversational aid [3]. This keyboard-input type conversation aid uses the keyboard to create sentences, which are then spoken by a voice synthesizer. It also has functions for registering frequently used words in categories and communicating using images. However, this keyboard input type conversation aid is difficult to use for severely ill ALS patients who have quadriplegia and have difficulty moving their bodies because the device requires the user to input data by hand. On the other hand, there is a gaze-input conversation aid [4]. Gaze input type conversation aids are devices that use a gaze input instead of a keyboard to create sentences and have them spoken by speech synthesis. Gaze-input conversation aids use a special gaze-input interface which is expensive and increases the system introduction cost. Common to these two products is the issue of not being able to converse with many people at once because they are face-to-face conversational types. As a previous study to solve these problems, there is an avatar robot developed by Oly Research Institute. This robot is installed in a place where ALS patients cannot go, and it can talk with a person whom the patient wants to talk with remotely by gestures. However, the robot needs to be set up at the place where there is the person whom the ALS patient wants to talk to, and the robot needs to be removed after the conversation is over, which requires the help of others besides the ALS patient. Another avatar technology other than robots is the Metaverse [6]. The avatars used in the metaverse are computer graphics, and users can select various characters of their choice. This technology allows people who have difficulty communicating in the real world to freely enjoy conversations and other activities in the virtual space through their favorite avatars. In this metaverse, there is no need to set up an avatar robot in advance at the place where you want to have a conversation, nor is it necessary to remove the avatar after the conversation is over. However, the interface currently used in the metaverse is a technology for able-bodied people, and it is difficult for people with physical disabilities, such as ALS patients, who cannot speak. On the other hand, a prior study of CG avatars is the Avatar Customer Service [5], which has begun to be introduced in convenience stores. This technology was developed to realize a new way of working that is not restricted by time, place, age, gender, or various other obstacles, and has been installed in actual convenience stores. The system displays a CG avatar on the display screen and uses gestures and hand gestures to converse with the customer. This technology does not burden others as avatar robots do because it is operated remotely using a personal computer, and it allows users to communicate remotely via avatars without questioning the location of the user. However, these prior technologies have issues that make it difficult to use in cases such as ALS patients with physical disabilities, because they are only for able-bodied people.

Therefore, in related work on the past, to the best of our knowledge, remote video communication for ALS patients with the help of eye-tracking does not exist. To solve this problem, we propose a remote video communication using eye- tracking for ALS patients who have difficulty going out. additionally, we propose a system that allows you to have a pleasant conversation with many people remotely using avatars, which are your favorite spoken dialogue agents.

3 PROPOSED SYSTEM

Section 3 provides a detailed overview of the avatar remote communication system for ALS patients using video conferencing tools proposed in this paper.

3.1 Outline of the Proposed System

Figure 1 shows a schematic diagram of the proposed system. The proposed system uses an online conferencing tool [2] as a telecommunication tool for the patient and the other party to have a conversation. There are two reasons for using an online conferencing tool in this paper. One is that ZOOM is widely used around the world as an online conferencing tool and is compatible with various platforms. The second reason is that ALS patients can easily communicate with their counterparts at home or in the hospital, regardless of their location, as long as the other party has an information terminal that can use the online conferencing tool and a network connection is available. We also thought that if the other party to the conversation used a smartphone or similar device, there would be no burden of removing the device after the conversation, and both parties could talk without burdening the other. Next, the proposed system uses an eye tracking interface [7] for the ALS patient's side. We adopted the Eye Tracking interface because it is less physically demanding than keyboard

Figure 1: ARCS-ALS overview

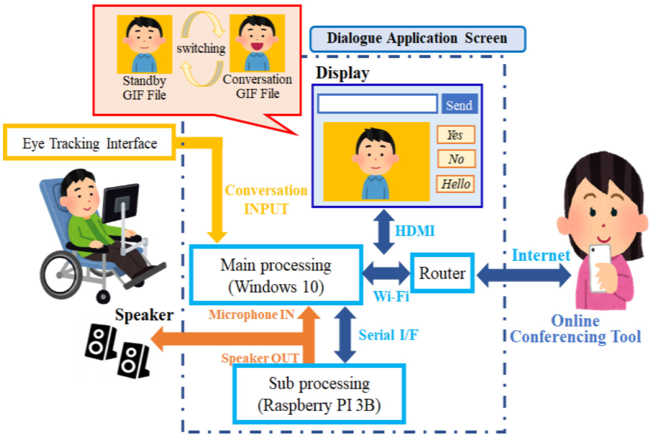


Figure 2: ARCS-ALS Prototype System Overview

operation, since severely ill ALS patients have significantly reduced physical abilities. In this system, the characters input by the eye tracking interface are directly output by text-to-speech to realize conversation. In addition, the system allows the user to freely add and select a character for the spoken dialogue agent that serves as the alter ego of the ALS patient by utilizing the function of adding and selecting spoken dialogue agents of a previous study [8]. As in the previous study, the ALS patient's favorite voice interaction agent can be used as an avatar by adding GIF (Graphics Interchange Format) files of his/her favorite cartoon character in conversation and in standby mode to the system, using FaceRig [9] or other software. The system allows users to use their favorite characters as avatars. By adding these two files to the system, a live-action video of the ALS patient himself can be used as a voice interaction agent. By adding a voice model for each character and configuring detailed settings such as voice, speed of speech, and endings, the system can output different voices for each character. The system allows users to remotely communicate with their conversation partners via the spoken dialogue agent by transmitting the display screen of the spoken dialogue agent selected by the user on the shared screen of an online conference tool. The system also has a branch speaker output so that the user can talk not only remotely but also with a person nearby. In this paper, we call the proposed system Avatar Remote Communication System for ALS Patients (ARCS-ALS).

4 PROTOTYPE DEVELOPMENT

Section 4 details the development of the ARCS-ALS prototype system proposed in Section 3.

4.1 Outline of Prototype System

Figure 2 shows an overview of the prototype system. In this prototype, distributed processing is performed by a Windows PC and a Raspberry PI3B+ (hereafter referred to as "RPI"). The Windows PC is in charge of the main processing, and the RPI is in charge of the sub-processing. The main processing on the Windows PC is the input processing of the Eye Tracking interface (model EyeTracke4C/Tobii) [7] used by the ALS patient to operate the ARCS-ALS, the switching processing of the video in standby and conversational mode according to the speech synthesis of the spoken dialogue agent that talks on behalf of the ALS patient, the input processing of the video in standby and conversational mode according to the speech synthesis of the spoken dialogue agent that talks on behalf of the ALS patient. The RPI is in charge of the sub-processing, and is in charge of the video switching between standby and conversation, as well as the remote communication with the other party via ZOOM. On the other hand, RPI, which is in charge of sub-processing, outputs the input from the Windows PC to the voice of the spoken dialogue agent and instructs the Windows PC when to switch between the live-action video in standby mode and the live-action video during conversation. The text-to-speech output from the sub-processing (RPI) was input to the microphone on the Windows PC side of the main processing via the speaker output of the RPI. This allows the synthesized voice to be output to the other party via ZOOM. In addition, the system is designed to output the synthesized voice from the sub-processing (RPI) to the speaker on the ALS patient's side through a branch. This enables the patient to talk not only online but also with people in the vicinity. Open JTalk [10] was used for the speech synthesis of the spoken dialogue agent.

4.2 Application Screen

In the prototype of the proposed system, ARCS-ALS, we developed a voice dialogue application to be used by ALS patients when conversing with each other. Figure 3 is a general view of the prototype system developed for this project, and Fig. 4 shows the screen of the application developed this time. As shown in Fig. 4, the large character in the middle of the screen is the selected character. In other words, it will be a voice dialogue agent that will carry out conversations on behalf of the ALS patient. When the patient wants to change the selected agent to suit his/her preference, he/she can click on the character display on the right side of the screen to change the character automatically.

In this prototype, we prepared four characters: the user's own image, a child character, and animal characters of a dog and a cat. Two methods were prepared for the input of

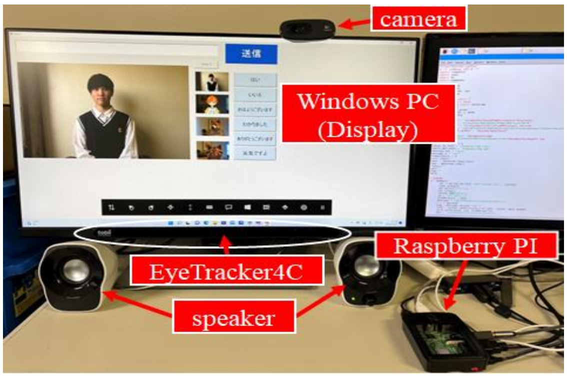


Figure 3: Prototype system developed for this project

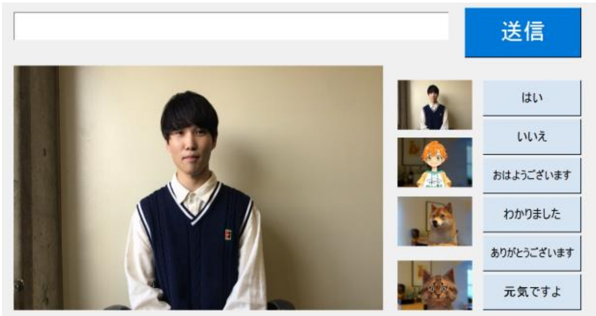


Figure 4: Application screen for ARCS-ALS

the conversation. In the first method, frequently used conversations such as replies, and greetings are registered in advance and can be spoken at the touch of a button. The second method is to use the keyboard, which is one of the functions of the eye control input interface, to input text. In this method, the ALS patient inputs what he/she wants to say using the keyboard and presses the "send" button to speak. Although it takes some time to input the content of the conversation, any kind of conversation can be sent.

4.3 Eye Tracking Interface

ARCS-ALS uses the Eye Tracking interface [7] as a method for ALS patients to talk. The reason for using this interface is that even ALS patients who have limb disabilities and have difficulty moving their bodies can use the Eye Tracking interface for eye control if they do not have eye problems. In our proposed system, we used an off-the-shelf eye tracking interface and the eye tracking function [7] that comes standard in Windows 10. To use the system, the Eye Tracking interface is connected to a Windows PC via a USB interface, and the main unit of the Eye Tracking interface is attached to the bottom of the monitor. After installing the driver, adjusting the position of the Eye Tracker 4C, and performing simple setups such as a visibility test, the system is able to detect the user's line of sight on the Windows PC screen. Next, by activating the eye control function that comes standard with Windows 10 and later, the user can click, use the mouse, and perform keyboard input, so even the physically challenged can operate the PC using only their line of sight. The Eye Tracking interface should be installed at an appropriate distance from the user, which is approximately 90 cm. If the

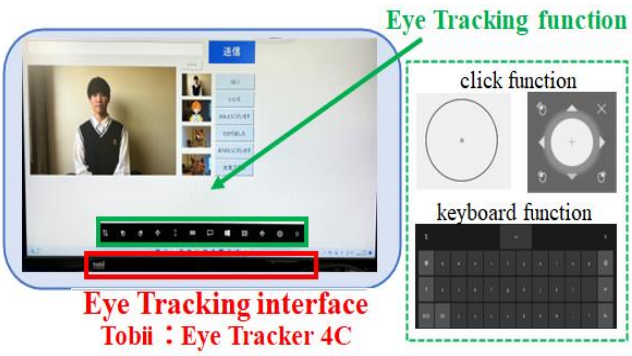


Figure 5: Overview of Eye Tracking interface

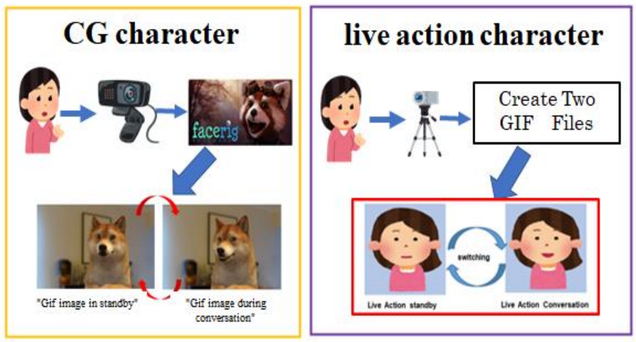


Figure 6 : How to Create Voice Dialogue Agents

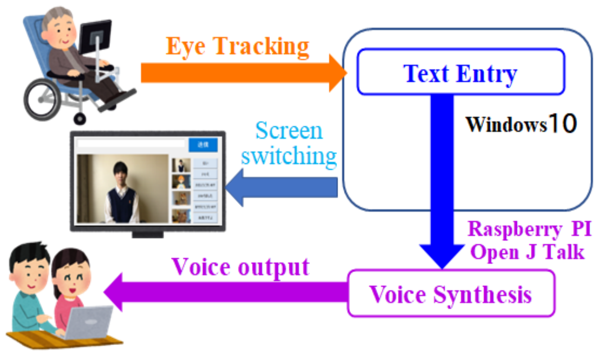


Figure 7: Speech synthesis output mechanism

distance is too close or too far, the camera cannot be calibrated accurately and may not operate properly.

Figure 5 shows the eye tracking interface used in the prototype and the eye control function. The eye tracking interface attached to the lower part of the monitor (circled in red) recognizes the user's line of sight, and the eye control function (circled in green) is used to control the operation. By looking at the desired function for a certain period of time, the user can instruct the eye tracking function to display icons to select the desired location for clicking and a keyboard for inputting text, as shown in Fig. 5. At first, it may be a little difficult to operate the function you want to use, but as you gradually become accustomed to it, you will be able to operate it at will.

4.4 Voice Dialogue Agent

In the ARCS-ALS system, a voice dialogue agent is used which talks to the other party instead of the patient. The voice dialogue agent was created by applying our previous studies. In this prototype, we prepared four characters: a character created from the ALS patient's own image, a child character, and animal characters (a dog and a cat). This allows the user to choose one of the characters according to his or her preference and environment, such as those who want to show their healthy self when talking to other persons or those who want to appear to be a substitute cartoon character instead of their own image. The voice dialogue agent can be created in two main ways: as a CG character or as a live-action character. Figure 6 shows a summary of each method. First, CG characters were created using FaceRig [9] software, which allows the user to become any character using a webcam, and two GIF images were created: one in the standby mode (as if the user is listening to a conversation) and one in the conversational mode (as if the user is talking). By displaying the two GIF images in accordance with the speech state of the voice dialogue agent using text-to-speech output and switching the character, the mouth movements are synchronized with the video as if the character is actually speaking. The live-action character is an application of our previous work on creating CG characters. To create a live-action character, two live-action videos were actually shot using video equipment, one in the standby mode and the other in the conversational mode.

As in the previous study, a live-action character is easily created by switching between the two GIF images in accordance with the speech state. In this case, it is necessary to shoot a short movie of 3 to 5 seconds and change the movie file to a GIF file. This makes it possible to repeat and play back the short video, which allows for longer text conversations. This method has the advantage that, once the user becomes accustomed to it, a live-action character can be created in about 10 minutes, making it possible to create a character of the person to be reproduced at a low cost. In addition, it is considered to be better to create the image of the ALS patient before the disease becomes severe, because it is considered to be practically difficult for the ALS patient to film himself after the disease has become severe Voice synthesis output.

The voice of the voice dialogue agent, which is the avatar of the ALS patient, is synthesized by Open J Talk [10], and its output is shown in Fig. 7. In this system, the ALS patient uses the eye control input interface to input the content of the conversation in text using the application developed in this study. The input conversation content is output as text-to-speech by Open J Talk [10], and the voice dialogue agent speaks the content. Acoustic models are prepared for each of the voice dialogue agents selected by the user, and the voice, speech style, and endings are changed for each character to speak. The dog CG character would end with "bow-wow" and the cat CG character would say "meow". On the other hand, when using a live-action character created from the ALS patient's own video, we prepared an acoustic model that closely resembled the patient's own voice, and then manually adjusted the parameters using Open J Talk to make the voice sound similar to the patient's own voice.

Specifically, the voice quality, which can be changed to feminine or masculine by changing related parameter values, the pitch shift, which changes the tone of the voice, and the speech speed, which changes the speed of speech, were adjusted to match the patient's voice. The values of these voice parameters were used because they can be fine-tuned, which actually takes time but produces highly reproducible voice. Other speech synthesis technologies include those that utilize AI technology to learn voices. Coestation [11] and LYREBIRD [12] are representative of these technologies. With these technologies, it is easy to produce a voice synthesizer that is close to the user's own voice by reading a few examples of sentences and having the AI learn those sentences. However, since the example sentences need to be learned repeatedly, it is not possible to train the AI for ALS patients who have difficulty in speech. Therefore, this system uses Open J Talk, which allows the user to manually adjust the parameters of the acoustic model based on the original voice.

5 EVALUATION

In this Section 5, a questionnaire evaluation was conducted to evaluate the prototype of the proposed system. This time, in addition to the questionnaire evaluation targeting ALS patients, the evaluation was conducted from three viewpoints targeting medical professionals who have knowledge about ALS patients, able-bodied persons. From the questionnaire evaluation, we investigated the effectiveness of the proposed system ARCS-ALS as a communication tool that enables ALS patients to enjoy talking with many people by their own power, and the functions and improvements that are required for this system.

5.1 Method of Evaluation

First, we conducted a questionnaire evaluation of one ALS patient through an acquaintance who had a close relationship with the patient. Face-to-face evaluation was difficult this time because of the coronary disaster. Although it was difficult to have the participants actually use the prototype, we asked them to understand the proposed system and to answer a questionnaire based on the assumption that they had used the prototype. This time, we evaluated the questionnaire on the assumption that the prototype was actually used, but we asked ALS patients who cooperated in this survey to fill out the questionnaire because they had experience using eye tracking interfaces in the past, and we thought they would give us their opinions. In the second evaluation, a questionnaire was sent to medical personnel involved with ALS patients. This time, a questionnaire evaluation was conducted on 26 medical personnel at several home health care nursing stations in Kanagawa Prefecture. In this survey, we created a video (about 9 minutes long) explaining the proposed system and showing an actual conversation using the prototype, so that the medical personnel could answer the questionnaire at any time according to their work schedule. The video was watched during breaks, etc., and the participants were asked to answer a questionnaire afterward. The Third evaluation was conducted with able-bodied persons. The questionnaire evaluation of able-bodied persons was conducted on 20 male subjects in their 20s. The survey method consisted of explaining the outline of the proposed system and the prototype to the subjects, having them watch a demonstration video of the actual use of the prototype (about 5 minutes), and then asking them for answering the questionnaire. The survey consisted of five major questions. These of evaluation procedure will be basically the same, just with different subjects.

5.2 Evaluation Results of ALS Patients

Table 1 shows the results of the questionnaire for ALS patients. The answer was "Yes" to the question 1 "Would you like to use the system you saw in the video? Yes" to the question "Would you like to use the system shown in the video? This indicates that the proposed system is useful from the viewpoint of ALS patients. In response to question 2, "In what ways would you like to use the system? The answers to all the questions were "I can use it by myself," "I can communicate easily," "I can communicate with others as my favorite character," and "I can talk with many people even if I cannot go out. In addition, "It is easy to listen to because of the voice synthesis and pronunciation intonation" and "Predictive text conversion is easy to use. These results indicate that the proposed system has achieved the objectives of the proposed system. Next, to question 3, "What would you like to use as a character for the spoken dialogue agent in this system? the answers were "My current live-action video," "My favorite animal (dog, cat, etc.) character," "My favorite animation," and so on. This suggests that the respondents would like to use not only their current live-action video, but also their favorite animals and favorite animation characters for conversation. In response to question 4, "What would you like to use as the voice of this system? the respondents answered "my voice","favorite star's voice", "favorite anime", etc., suggesting that they would like to select a voice that matches the avatar character they are using. Next, There were no responses to the question 5, "What improvements do you think are needed in the layout of the application screen of the system? In response to question 6, "In addition to the remote conversation function, what other service functions would you like to see in this system? was "more types of keyboards" as other opinions. Next, There was no answer to the question 7, "What do you think needs to be improved in the future? Finally, to the question 8, "Please feel free to give us any other comments or requests you may have. The last question was "Please feel free to give us any other comments or requests. I use the internet, YouTube, Facebook, LINE, and recently google spreadsheets instead of Excel. The rest I rarely use, so I thought it would be better to create a well-balanced system. Please let me make a decision after actually using the system. In summary, for ALS patients who responded to this survey, in addition to the avatar remote communication function, it would be useful to have various functions that would help them in their daily lives, such as remote control of home appliances, health management, and a doorbell.

Table1:Results of Questionnaire Evaluation of ALS Patient.

question	choice	Answer
Q1: Would you like to actually use the system you have seen in this video? (Or would you like to use the system you saw in the video?)	1. yes 2. no	1
Q2. If you answered "yes" to Q1. In what ways would you like to use it?	1. because I can use it by myself 2. because it is easy to communicate 3. because I can communicate with others as my favorite character 4. because I can communicate with many people even if I cannot go out 5. other	1 2 3 4 5 [It is easy to listen to the voice synthesizer, and the pronunciation can be inflected. Predictive conversion is also easy to use.]
Q3. What would you like to use as a character for the spoken dialogue agent in this system? Please select more than one.	1. past live-action videos of yourself when you were healthy 2. live-action video of yourself in the present 3. favorite animal (dog, cat, etc.) character 4. favorite cartoon or other character 5. other	2 3 4 5 [Furniture, musical instruments, and food]
Q4. what would you like to use as the voice of this system? Please select more than one.	1. your own voice 2. voice of your favorite star 3. voice of your favorite cartoon character 4. other	1 2 3
Q5. What improvements do you think are needed in the layout of the application screen of the system? Please select more than one that apply.	1. difficult to select canned text 2. the text box is too small and it is hard to input characters 3. difficult to select characters 4. small screen size 5. small character size 6. other	
Q6. In addition to the remote conversation function, what other service functions would you like to see in this system? Please select more than one that apply.	1. subtitling of conversations 2. function to remotely control home appliances 3. function to measure and manage the user's health (body temperature, pulse wave, Spo2, etc.) 4. function to provide daily life information such as weather, time, etc. 5. other 5. other	1 2 3 4 5 [More doorbells, zoom function, and keyboard types]
Q7. What do you think needs to be improved in the future? Please select more than one.	1. speed of conversation 2. awkwardness of response 3. gaze input 4. application 5. screen size 6. audio quality 7. other	
Q8. Please feel free to give us any other comments or requests.	I think it is better to start with less, as more features will cost more. I use the internet, YouTube, Facebook, LINE, and recently google spreadsheets instead of Excel. I rarely use the rest, so I thought it would be better to create a well-balanced system. Please let me judge the rest by actually using it.	

but however, considering the cost at the time of installation, we feel that a minimal number of functions is sufficient. In other words, in addition to the proposed remote communication function, it would be good if the user could watch YouTube videos on the Internet or send messages via SNS.

5.3 Evaluation Results of Medical Personnel

Figure 8 shows the results of the questionnaire for healthcare professionals involved with ALS patients. This time, the questionnaire was sent to 26 medical professionals involved with ALS patients. The gender of the respondents was 22 females and 4 males, and the age breakdown was as follows: 1 respondent was in her 20s, 8 were in their 30s, 10 were in their 40s, 5 were in their 50s, and 2 were in their 60s or older. To the question 1, "Do you feel that the system you saw in the demo video is a system that you would recommend to ALS patients? 96% of all respondents answered "Yes" and 4% answered "No". This indicates that more than 90% of the medical professionals highly evaluated the proposed system.

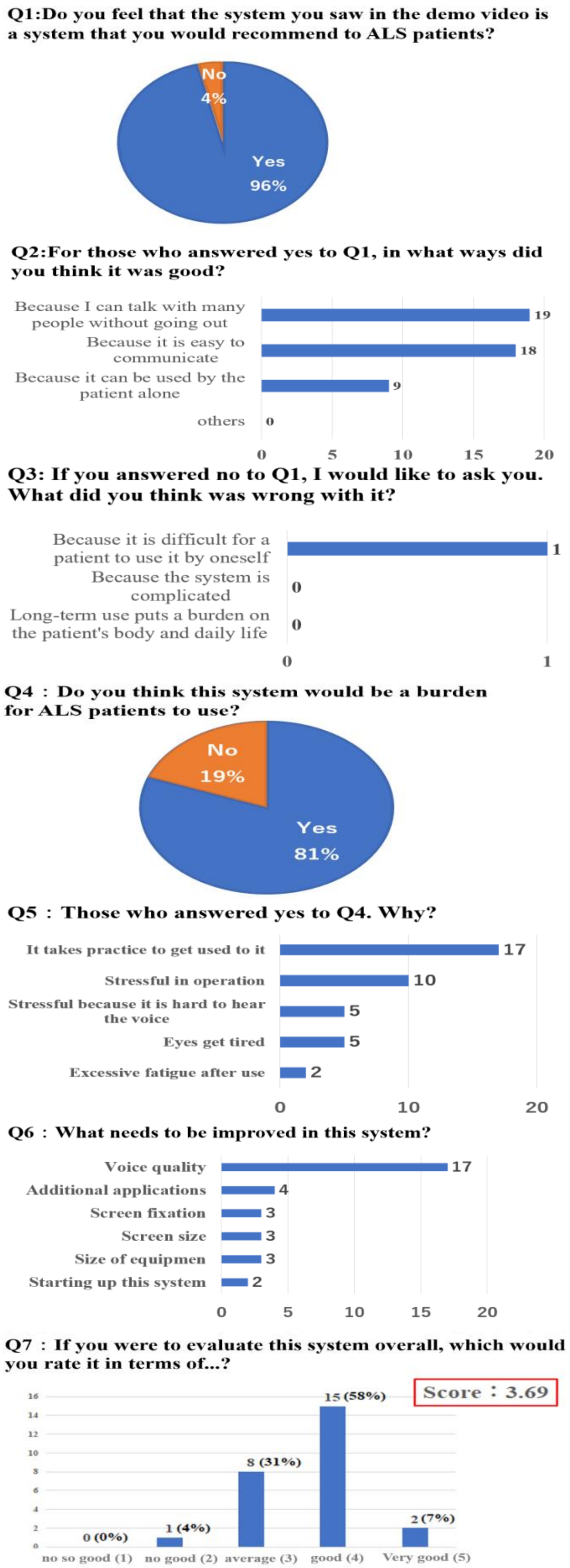


Figure 8: Questionnaire evaluation results for medical professionals.

Next, in response to question 2, "For those who answered yes to question 1, in what ways did you think it was good? " The most common answer was "Because I can talk with many people without going out" was 19 responses, followed by "Because it is easy to communicate" was 18 responses, and then "Because it can be used by the patient alone" was 9 responses. These results indicate that the proposed system's purpose of communicating with many people without having to go outside and its ease of use were well understood and appreciated by many medical professionals. Next, in response to question 3 of "If you answered no to question 1, I would like to ask you. What did you think was wrong with it?" one respondent answered "Because it is difficult for one patient to use it alone. Next, in response to question 4 of "Do you think this system would be a burden for ALS patients to use?" about 81% of all respondents answered "Yes," and about 19% of all respondents answered "No". In response to Question 5 of " Those who answered yes to question 4. Why?" , the most common answer was "It takes practice to get used to it" was 17 responses, "Stressful in operation" was the second most common was 10 responses, and "Stressful because it is hard to hear the voice" and "Eyes get tired" were the third most common was 5 responses. This indicates that many respondents thought that the system requires practice before they get used to it, and that many healthcare professionals thought that the system would be stressful to operate. Next, in response to question 6, "What needs to be improved in this system?" the most common response was "voice quality" with 17 responses.

Other responses were "additional applications" was 4 responses, "screen fixation," "size of equipment," and "screen size" was 3 responses, and "starting up this system" was 2 responses. This suggests that the current system is difficult for ALS patients to listen to due to the poor voice quality of the avatar character speaking. Finally, to question 7, "If you were to evaluate this system overall, which would you rate it in terms of...?" 7% of all respondents answered "very good (5)," 58% answered "good (4)," 31% answered "average (3)," 4% answered "not good (2)," and 0% answered "not so good (1)." The total of "very good" and "good" responses amounted to 65% overall. The average score on a 5-point scale from "very good: 5" to "not so good: 1" was about 3.69, which is a reasonably high rating. Therefore, although the proposed system needs to be improved in terms of "voice quality" and "requires practice until the user becomes familiar with the operation," the system is considered to be reasonably useful for ALS patients from the viewpoint of medical professionals who are usually involved with ALS patients.

5.4 Evaluation Results of Able-bodied Persons

Figure 9 shows the results of the questionnaire with able-bodied persons. All respondents answered "Yes" to the first question, "Do you feel that you would recommend this system to ALS patients?" For those who answered "yes," the second question was asked: "In what ways do you think this system is recommended for ALS patients?”

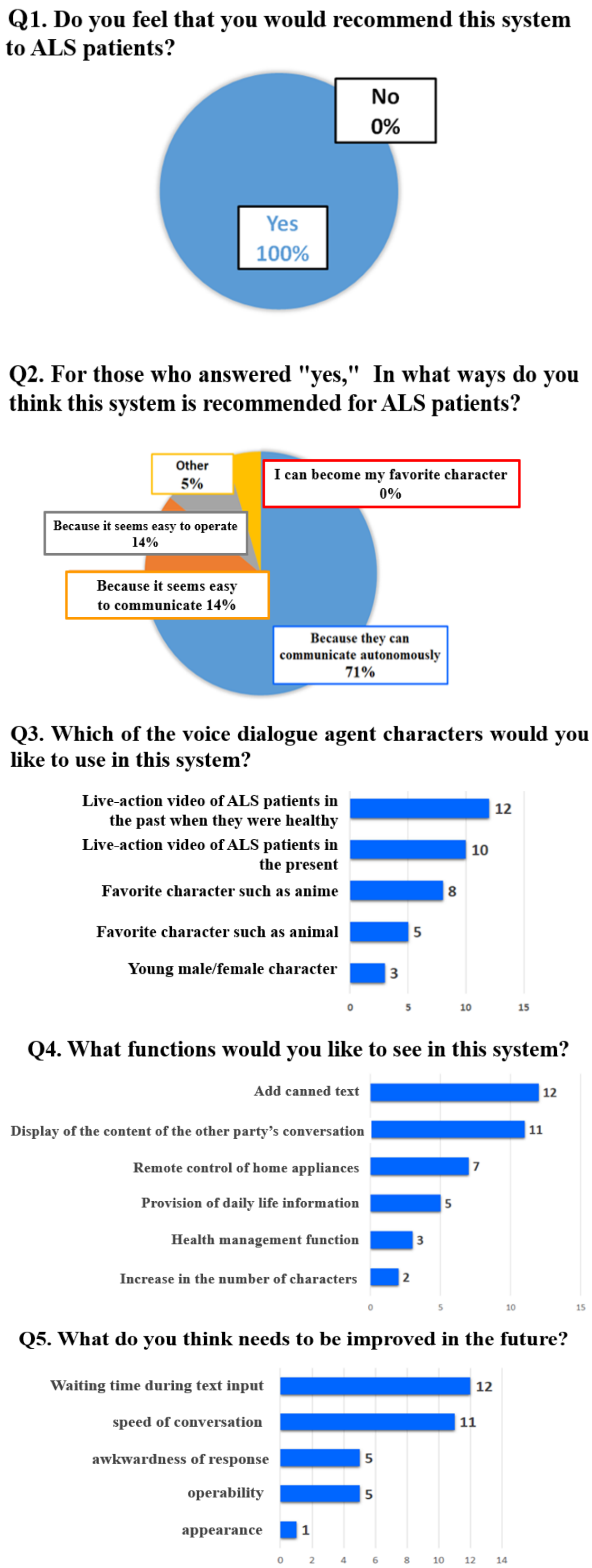


Figure 9: Questionnaire evaluation results for able-bodied persons.

The most popular answer was “Because they can communicate autonomously (71%)”. The next most common answer was "Because it seems easy to communicate" (14%), followed by "Because it seems easy to operate" (10%), and "Because I can become my favorite character" (0%). The other responses were "the cost of introduction is low" and "it is easy to use”. Next, Question 3 was asked: "Which of the voice dialogue agent characters would you like to use in this system?” The most common answer was "Live-action video of ALS patients in the past when they were healthy" (12 responses), and "Live-action video of ALS patients in the present" (10 responses), both of which were answered by more than one half of all the respondents. Among the other responses, "Favorite character such as anime" was selected by 8 respondents, "Favorite character such as animal" by 5 respondents, and "Young male/female character" by 3 respondents, indicating that "Favorite character such as anime" was the most sought-after character for ALS patients other than live-action videos. As for Question 4, "What functions would you like to see in this system?”, the two most popular responses were “add canned text” (12) and “display of the content of the other party’s conversation” (11). The next most common responses were "remote control of home appliances" (7), "provision of daily life information" (5), "health management function" (3), and "increase in the number of characters" (2). Finally, to Question 5, “What do you think needs to be improved in the future?”, the most common response was "waiting time during text input" (12 responses), followed by "speed of conversation" (11 responses), both of which were cited by more than one half of the respondents as an area that needs improvement. Next were "awkwardness of response" and "operability" (5 responses each), and "appearance" (1 response).

5.5 Summary of Evaluation Results

In this prototype evaluation, the proposed system was evaluated from various perspectives by ALS patients, medical professionals working with ALS patients and able-bodied persons. As a result, more than 90% of the respondents answered that they would recommend the proposed system to ALS patients, and the system was highly evaluated. The reasons for this are as follows: "I can use the system by myself," "I can communicate easily," "I can communicate with others as my favorite character," and "I can talk with many people even if I cannot go out," according to the questionnaire results from the ALS patients' side.

On the other hand, the most common response from medical professionals involved with ALS patients in the questionnaire was "I like the fact that I can talk with many people without going out", followed by "I can communicate easily" and "The patient can use the system alone". This indicates that the medical professionals involved with ALS patients understood and highly evaluated the features of the proposed system. On the other hand, however, when medical professionals were asked, "Do you think this system will be a burden for ALS patients? 81% of the respondents answered that the system would be a burden for ALS patients. The reason for this is that it requires practice until they become accustomed to its operation. Also, some respondents felt that the proposed system is burdensome in terms of operation. This point is considered to be an issue for the future of this system. Next, we asked, "Are there any functions you would like to see in the proposed system? The ALS patient's answers were "a function to display conversations," "a function to remotely operate home appliances," "a function to measure and manage the user's health (temperature, pulse, SpO2)," "a function to provide daily life information such as weather and time," etc., as well as "a doorbell and zoom function.

However, in consideration of the cost at the time of installation, ALS patients' opinions indicated that, as a minimum, in addition to the proposed remote avatar communication function, they would like to be able to watch YouTube videos on the Internet and send messages via SNS. Furthermore, when asked, "What do you think needs to be improved in the future?" the ALS patients answered that there were no areas that needed improvement, while the medical professionals most frequently asked for improvement in "voice quality. This indicates that the quality of the voice output from the avatar is something that needs to be improved in the future. In addition, the medical professionals who responded to the questionnaire gave the proposed system a five-point overall rating, with an average score of 3.69, indicating that they gave the system a reasonably high rating.

6 CONCLUSION

In this paper, we proposed a remote video communication using eye-tracking for ALS patients who have difficulty going out. additionally, we proposed a system that allows you to have a pleasant conversation with many people remotely using avatars, which are your favorite spoken dialogue agents. The proposed system utilizes an online conferencing tool as a method of conversation with other parties, and uses an interface that can be operated by the ALS patient himself/herself by using an eye tracking interface. In addition, we have developed a new dedicated application for ALS patients to conduct a conversation. This enables ALS patients to easily converse with many people at their own will without the help of others. In addition, we developed an actual prototype of the proposed system and evaluated the prototype. In the evaluation, the prototype was evaluated from two perspectives: a questionnaire evaluation for ALS patients themselves and a questionnaire evaluation for medical professionals involved with ALS patients. The evaluation results confirmed the effectiveness of the proposed system, ARCS-ALS. We were also able to confirm the required functions of the system and points to be improved.

7 FUTURE WORK

Through the development and evaluation of the prototype of the proposed system, we have learned some details that need to be improved, and we plan to improve this prototype for practical use in the future.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to the TOWN NURSE Home-visit Nurse Station, KAWASAKI Home-visit Nursing Station, HIBARINOMORI Home-visit Nurse Station, and KAWASAKIK DAISHID Home-visit Nurse Station for their cooperation in this study. and We would like to express our sincere gratitude to Mr. Riku Sugimoto and Mr. Shouma Hamada for their cooperation in this study.

This study was approved by the Ethical Review Board for the use of human subjects of Kanagawa Institute of Technology (No.20220715-01).

REFERENCES

[1] "Orihime Product Introduction ", < https://orihime.orylab.com/>, (Referred May.2022).

[2] “Zoom”, < https://zoom.us/jp-jp/meetings.html>, (Referred May.2022).

[3] “TALKING AID+”, <https://www.talkingaid.net/products/ta-plus>, (Referred May.2022).

[4] "LUCY Double Giken ", <http://www.j-d.co.jp/fukushikiki-lucy.html>, (Referred May.2022).

[5] "LawsonLaboratory", <https://www.lawson.co.jp/lab/tsuushin/art/1458363_4659.html>,(Referred May. 2022).

[6] "Metaverse Japan", <https://metaverse-japan.org/>, (Referred May.2022).

[7] Eye Tracker Product List|Tobii Technology <https://www.tobiipro.com/ja/product-listing>, (Referred May.2022).

[8] N. Kokubu, T.Mukai, and K.Abe, " A Proposal for a Low-Cost voice Dialogue System Using Live-Action Video Contents ", 2021 IEEE Global Conference on Consumer Electronics Proceedings (GCCE2021), pp.340-341, Oct. (2021).

[9] FaceRig Live2D Module, <https://www.live2d.com/in-terviews/facerig>, (Referred May.2022).

[10] Open J Talk, < http://open-jtalk.sp.nitech.ac.jp>, (Referred May.2022).

[11] "coestation", <https://coestation.jp/consumer/coestationapp/>, (Referred May.2022).

[12] "Lyrebird", <https://www.descript.com/lyrebird>, (Referred May.2022).

(Received: December 8, 2022)

(Accepted: October 4, 2023)

Nagisa Kokubu received the M. E. degree from Kanagawa Institute of Technology in 2023. His research interests include Avatar Remote Communication System for ALS patients, Voice Dialogue System. He is currently an sales engineer in TOKYO ELECTRIC POWER COMPANY HOLDINGS. INC.

Takayo Namba received a Ph.D. from Nagano College of Nursing in 2009. She was a professor of the faculty of nursing, Kanagawa Institute of Technology from Apr.2017 to Mar.2019. She was a professor of Teikyo Heisei University from Apr.2021 to Mar.2022. She has been a professor of Kawasaki City College of Nursing since Apr.2022. Her research interests about Elderly Abuse and Co-dependency, Educational program and Visiting nurses. She is a member of the Japan Academy of Nursing Science and the Japan Society for Health and Welfare etc.

Keiichi Abe received a Ph. D. in informatics from Shizuoka University in 2012. He has been a professor of the Faculty of Creative Engineering, Kanagawa Institute of Technology. His research interests include smart system using IOT technology, Interactive system using Robot and CG, and Victims Information Management System At large scale disaster, HEMS. He is a member of IPSJ, IEICE, ACM and IEEE.

Submission Guidance

About IJIS

International Journal of Informatics Society (ISSN 1883-4566) is published in one volume of three issues a year. One should be a member of Informatics Society for the submission of the article at least. A submission article is reviewed at least two reviewer. The online version of the journal is available at the following site: http://www.infsoc.org.

Aims and Scope of Informatics Society

The evolution of informatics heralds a new information society. It provides more convenience to our life. Informatics and technologies have been integrated by various fields. For example, mathematics, linguistics, logics, engineering, and new fields will join it. Especially, we are continuing to maintain an awareness of informatics and communication convergence. Informatics Society is the organization that tries to develop informatics and technologies with this convergence. International Journal of Informatics Society (IJIS) is the journal of Informatics Society.

Areas of interest include, but are not limited to:

Internet of Things (IoT)
Smart Cities, Communities, and Spaces
Big Data, Artificial Intelligence, and Data Science
Network Systems and Protocols
Computer Supported Cooperative Work and Groupware
Security and Privacy in Information Systems
Intelligent Transportation System
Distributed Computing
Multi-media communication
Information systems
Mobile computing
Ubiquitous computing

Instruction to Authors

For detailed instructions please refer to the Authors Corner on our Web site, http://www.infsoc.org/.

Submission of manuscripts: There is no limitation of page count as full papers, each of which will be subject to a full review process. An electronic, PDF-based submission of papers is mandatory. Download and use the LaTeX2e or Microsoft Word sample IJIS formats.

http://www.infsoc.org/IJIS-Format.pdf

LaTeX2e

LaTeX2e files (ZIP) http://www.infsoc.org/template_IJIS.zip

Microsoft Word™

Sample document http://www.infsoc.org/sample_IJIS.doc

Please send the PDF file of your paper to secretariat@infsoc.org with the following information:

Title, Author: Name (Affiliation), Name (Affiliation), Corresponding Author. Address, Tel, Fax, E-mail:

Copyright

Publisher

Address: Informatics Laboratory, 3-41 Tsujimachi, Kitaku, Nagoya 462-0032, Japan
E-mail: secretariat@infsoc.org

CONTENTS