International Journal of Service and Knowledge Management
International Institute of Applied Informatics
2023, Vol. 7, No. 1, IJSKM733

Evaluations of the Service Robot OSONO Referring to Joruri Puppets and a Remote Evaluation System with Operation and Directing Elements

Masahiko Narita [*], Toru Izui [†], Sachiko Nakagawa [‡], Yasufumi Takama [§], Nobuto Matsuhira [**]

Abstract

One of the important requirements for service robots is attracting people by attractiveness and to be able to exchange messages with people. With reference to the traditional Japanese puppet show, Ningyo Joruri, we have independently developed OSONO, which is a physical robot, with high-quality choreography, we have exhibited OSONO through large-scale exhibitions and evaluated it. In this paper, we report questionnaire evaluations on OSONO targeting a puppeteer expert group and compare them with existing results targeting the ordinary person. This shows that the method of creating OSONO and its choreography is effective in widely general. Additionally, it is verified that the same result can be obtained in the remote evaluation with the robot operation. Moreover, we evaluate the effect by adding the technique of Ningyo- Joruri to the operation and sound effects of OSONO, and analyze the influence of the quality of the real-time image transferring quality. Through these results, we propose effective system design guidelines for displaying robots online.

Keywords: Robot Services, Active Sensing, Physical Properties, Choreography, Service Robot, Remote evaluation system, Joruri Puppet, RSNP (Robot Service Network Protocol)

1 Introduction

In the society where a wide variety of service robots are used in cooperation, the Common Robot User Interface that eliminates inconveniences due to differences in operations depending on models and manufacturers is required. Therefore, it is considered that one of the important requirements for service robots is to attract people and to be able to exchange messages with peo-

[*] Advanced Institute of Graduate School of Industrial Technology (AIIT) , Tokyo, Japan
[†] EvolveITWorks Corporation, Tokyo, Japan
[‡] Aoyama Gakuin University, Tokyo, Japan
[§] Tokyo Metoropolitan University, Tokyo, Japan
[**] Shibaura Institute of Technology Tokyo, Japan

ple smoothly. In order to realize this, we are conducting research on gestures and choreography. In addition, with reference to the traditional Japanese puppet show "Ningyo Joruri," we have independently been developing a robot named OSONO, which has the physicality and high-quality choreography, and are evaluating OSONO in several cases [1][2].

The rest of this paper is organized as follows. In Chapter 2, we briefly describe an overview of the robot OSONO and also describe an overview of the questionnaire results by the ordinary person that conducted in the past. In Chapter 3, we report questionnaire evaluations by a puppeteer expert group and compare them with the conventional evaluation by the ordinary person. These results are feedbacked to the robot. In Chapter 4 and Chapter 5, we design and develop two remote evaluation systems, a basic version for remote evaluation with operation and an enhanced version that improves the real-time image transferring quality and adds directing elements of Ningyo Joruri. In Chapter 6, we carry out evaluations using them. In the evaluation, we verify the evaluation result is the same as the local (face-to-face) evaluation, and the directing elements of Ningyo Joruri to the operation and sound effects, and analyze the influence of the quality of the real-time image transferring quality. Through these results, we propose effective system design guidelines for displaying robots in online events.

2 Overview of Robot "OSONO "

2.1 Physicality of the Robot

It is known that robots with physicality are more interesting to visitors and users than systems that transmit information through displays [3]. Physicality is evaluated not only by size but also by quality [4][5]. Since it changes greatly depending on the situation, the problem is the methodology for constructing a robot that can attract more people and exchange business messages.

Figure 1: OSONO in "robot showcase" online

We aim for effective expression by incorporating the works inherited over time. To begin with, we developed a Kashira-robot (a head-only robot) using a unique design method with reference to Ningyo Joruri. Then we developed more a robot with physicality, named OSONO. With the cooperation of the national important intangible cultural property holding organization "Shimonaka-za", OSONO's choreography has been realized by linear approximation of feature points, which were extracted from 30 seconds of the opening of Ningyo Joruri's famous scene "OSONO no KUDOKI (OSONO’s saying)” of “SAKAYA no Dan (The scene of a liquor store)”. AI technology OpenPose [6] was used for feature points extraction. Finally, we obtained the choreography in good quality [1].
OSONO was exhibited in several events. For example, at the International Robot Exhibition 2019 in Tokyo, where many engineers, children, young women, etc. stopped in front of OSONO to look at the robot, take pictures, react to the robot itself, react to its choreography, and so on. It was well-received. Moreover, using our remote evaluation system described in Chapter 4, we demonstrated OSONO at “robot showcase” online exhibition hosted by the Ministry of Education, Culture, Sports, Science and Technology of JAPAN Government in 2021 [7] (Figure 1).

2.2 Short Introduction of Ningyo Joruri

Ningyo-Joruri is “Japanese puppet show” registered as one of World Heritage. It was very popular during 1700-1900. Ningyo-Joruri is called “Bunraku”, sometimes. Bunraku is still performed on a commercial basis. On the other hand, Ningyo-Joruri was performed as an entertainment in post towns in the old days. There are still some groups to remain, and their skill have been handed down. "Shimonaka-za" described in the previous section is one of such traditional groups. Ningyo-Joruri is played by the combination of the narration and the lines/quote by “Tayu”, Music by “Shamisen” (string sound), and Joruri-puppets with 3 puppeteers for one puppet.

2.3 Validity Verification of OSONO by the Ordinary Person

To verify the effectiveness, a questionnaire evaluation by the ordinary person was conducted to

Figure 2: Puppteer exerts evaluating OSONO

evaluate (1) OSONO's attractiveness, (2) choreography quality, (3) how few actuators can realize the choreography without impairing quality, (4) small effects such as eye blinking and lip-synching [1]. The evaluation was carried out for three groups, 32 evaluators in total (the ordinary person: master students in information processing, faculty members, and robot researchers). The evaluation objects were the actual robot and simulation of the robot OSONO, which had movable eyes and mouth and movable actuators for the body and neck from 2 to 4. As the questionnaire, multiple robots OSONO with different numbers of actuators, which are evaluation objects, present the choreography to the evaluators. They answer the questions. As a result, many of them find that the actual robot is attractive and the quality of the choreography is good. In addition, they answered that three actuators case is the best and does not deteriorate the choreography quality. So, it was found that in the case of this choreography, many actuators are not always required to achieve high-quality choreography.

3 Evaluation by Expert Puppeteer

We have been developing and evaluating the robot with referring to Joruri Puppet. Therefore, it is very interested in whether the evaluation result is different for puppeteers, who have a lot of knowledge and experience on Joruri Puppet and operate them frequently. In this chapter, we report the results that we conduct the questionnaire equivalent to that for the ordinary person mentioned in Section 2.3 to the puppeteer expert group, compare the evaluation results of both, consider their comments, and try to reflect them on the robot.

3.1 Questionnaire Evaluation

The puppeteer expert group consists of 17 people in total, from 20 to 70 years old and 2 to 30 years of experience in the Sagami-ningyo-shibai “Shimonaka-za” mentioned above. The evaluators answer the questionnaire while watching the choreography of actual moving OSONO robots. The robots to be evaluated were shown in Figure 1. Figure 2 shows the scene they are conducting the questionnaire. The movable actuators of the robot are the body (Rotation, Tilt forward/backward, Tilt left/ right) and neck (Tilt front/ back). The test cases A, B, C, and D in the tables correspond to the number of movable actuators 4, 3, 2 (neck’s Tilt front/ back move was fixed) and 2 (body’s Tilt forward/ backward). Actual questions in the questionnaire of Table 1 and Table 2 are: (1) OSONO is attractive as a robot, (2-1) No discomfort in the choreography, (2-2) OSONO’s choreography is attractive. The results of the questionnaire are shown in Table 1 and Table 3.

Please note that in order to facilitate comparison with the experimental results so far, this paper uses the score of the questionnaire to be 1 to 5, which means that the lower the score is, the higher the evaluation is.

Table 1: Evaluation results by actual robot (Puppeteer expert)

Test Case (described in 3)	A	B	D
(1) Attractive	2.4	2.2	2.4
(2-1) No discomfort in the choreography	3.3	3.3	3.3
(2-2) Choreography is easy to understand	2.6	2.8	2.9
Average of above 3 items	2.8	2.8	2.9

Note: lower the score, the higher the evaluation in this paper.

Table 2: Evaluation results by actual robot (Ordinary person) [8]

Test Case	B	C	D
(1) Attractive	2.1	2.1	2.3
(2-1) No discomfort in the choreography	2.8	2.9	2.7
(2-2) Choreography is easy to understand	2.5	2.5	2.7
Average of above 3 items	2.5	2.5	2.6

Table 3 Evaluation results for Effects: the lower the score is, the higher the evaluation is.

	Puppeteer experts	Ordinary persons [8]
(4-1) Eye blinking is effective?	2.5	1.8
(4-2) Opening/closing the mouth is effective?	2.9	2.3
(4-3) Fine head nodding is effective?	2.3	2.2

Table 2 shows the results of the ordinary person obtained in 2.3 corresponding to Table 1 for comparison. Actual questions in the questionnaire of Table 3 are: (4-1) Eye blinking is effective? (4-2) Opening/closing the mouth is effective? (4-3) Fine head nodding is effective? Note: each score is 1 to 5, and a lower score means a better evaluation. Table 3 shows the evaluation results on small effects such as eye blinking and lip-synching by the puppeteer expert group, the same questions for the ordinary person mentioned in Section 2.3. The results of the questionnaire are a little bit worse overall among the puppeteers. This is probably because their natural blinking and frequent lip-synching are not found in conventional Ningyo Joruri performances. However, since the scores of "attractive" for Chapter 2 (three actuators) are 2.1 or 2.2, they both see OSONO as attractive. Other Details will be described below.

· Regarding on “No discomfort in the choreography”, puppeteers react sensitively than the ordinary person.

· On the other hand, regarding the effects of eye blink and mouth opening/closing that Ningyo Joruri does not have, the ordinary person is positive on them, but about half of the puppeteers have the disagreement, and as the result, the average value has dropped.

· In the free form comments, 8 out of 17 people pointed out “jerkiness" in the choreography, and advised that it is better to move faster or slower. Additionally, the puppeteers, who have 25 years and 23 years experiences, want youthfulness on OSONO’s head. OSONO's face color is beige close to pink, however women heads of Joruri Puppet are pure white in general, so that there are negative reactions to the face color.

3.2 Reflection of Comments

We try to resolve two of the most important comments from the evaluators in the previous section.

· Unsmooth move

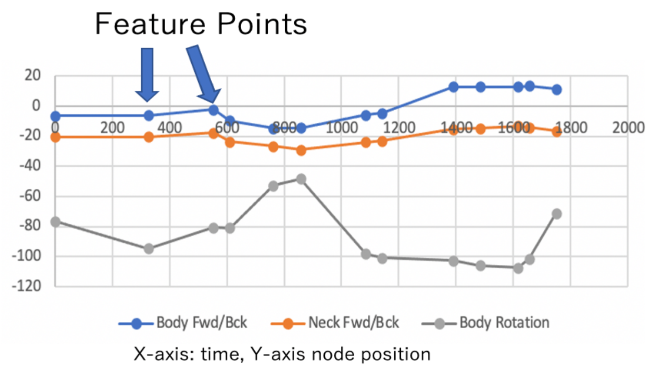

Figure 3: Features points and Operation

Unsmooth movement is caused by low actuator resolution. For example, when the head moves down, it is finely braked, resulting non-smooth motion occurred. Therefore, we chose an important part on the performance to eliminate un-smoothness. We used "overshoot" method, which is one of the methods to emphasize the choreography, that starts by moving in the opposite direction to the target direction and then moves in the target direction (Figure 3). At the same time, the movement speed can be 4.2 times up without making the choreography unnatural and the swing total time did not change. Finally, the jerkiness was decreased about 20% when measuring 3-axis accelerometer attached. As a result, the choreography became larger, clearer and better.

· Youthfulness

There are countless ways to emphasize youthfulness, we chose to make the eyebrows thicker and longer, shade the inner and outer corners of the eyes with brown. We also made the chin slightly higher, and the bottoms of both cheeks rounded and spread. In addition, by thinning the nasolabial fold slightly forward on the outside of the chin and mouth, the problem was solved with a thicker cheek (Figure 4).

3.3 Summary

In the evaluation of test case of B by the puppeteer expert group on Table 2, the score of "Attractive" is 2.2. Accordingly, OSONO is evaluated as "Attractive" as in the ordinary person case, and the target choreography can be well expressed by three actuators. This shows that the method of creating OSONO and its choreography is effective in widely general.

· The puppeteer responds sensitively to a sense of incongruity of the choreography, evaluated score is 3.3,“ as the same as the ordinary persons.

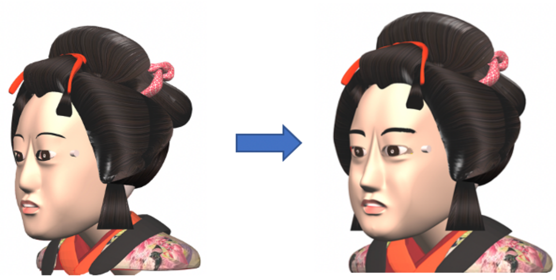

Figure 4: A model that emphasizes youthfulness (Right)

· About the sense of incongruity and youth, which were pointed out by evaluators, were resolved by modifying the choreography and remodeling the head.

These evaluations suggest that the choreography approach with reference to the Joruri puppet works well. We would like to expand more choreographies and to apply them to service robots.

4 Basic Remote Evaluation System

The remote evaluation is necessary when the face to face evaluation is difficult to carry out, and useful when many evaluators are needed to conduct the evaluation, In this chapter, in order to carry out a remote evaluation, we point out the basic requirements for a remote evaluation system. And we design and development of a remote evaluation system with Open Source software to satisfy them and technical issues [9].

4.1 Requirements of the Remote Evaluation System

The purpose of this system is to realize a remote evaluation system that replaces the conventional evaluation method, which is performed face-to-face, with the evaluation method that the evaluator observes the autonomous movement of the robot, answers the questionnaire and the answers are aggregated. Specifically, (1) the evaluator can remotely observe the robot performing various choreography, (2) evaluation results can be obtained immediately through questionnaires, (3) the evaluator can operate the robot's choreography with acceptable delay, and (4) it is possible to carry out more evaluators simultaneously than face to face questionnaire. In addition, (5) extensibility that can evaluate the movement of multiple robots in cooperation and the reaction of robots in conversation with humans, in future. (6) In order to exhibit at an online event and carry out an evaluation, it is also necessary to create a function to link with an online event.

In order to realize these requirements, we constructed the following remote evaluation system shown in Figure 5 (red part is for the extended remote system described in Chapter 5): (1) The robot to be evaluated and the evaluator's PC/ smartphone are connected via a server, (2) The operation of the robot is autonomously or remotely controlled by the evaluator via the server with buttons and so on, (3) The movement of the robot is captured and image-distributed to the evaluator's terminal through the camera on the robot side and the cache of the server. (4) Information on robot's operations and evaluators are stored as a log on the server. (5) Online questionnaires are set up on PCs and smartphones so that evaluators can easily answer the ques-

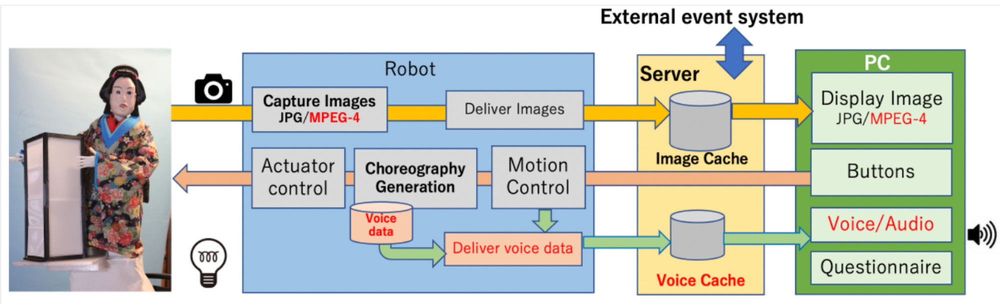

Figure 5: Outline of Basic and Extended Remote Evaluation System
(Red parts is for the extended remote evaluation system)

tionnaires. In addition, the information collected in (4) and (5) will be subject to later evaluation and analysis. (6) Cooperate with other online events as needed.

The purpose of implementing this system is to remotely achieve the same evaluation as face-to-face evaluation, and the problem is to ensure the reality. Technically, it is important to smoothly deliver high-quality, real-time video of robot movements and achieve remote control within an acceptable time delay. The points to consider in the implementation are focused on solving the problem. As described later, if you get enough reality, you will have a unique sense of reality that you had become a Puppeteer, and our remote evaluation system will function like a remote acting system.

4.2 Implementation of the Remote Evaluation System

In this section, we describe the implementation of the remote evaluation system. The main part uses the non-professional robot service platform [10] and RSNP (Robot Service Network Protocol) [11][12], as its base, which we have used in many robot service research projects. RSNP has the function of transferring images from the robot, named DistributeCameraImage, and operating the robot, named Motion profile, from the server, and Java-based toolkits and sample code are provided. The communication between the server and PC/Smart Phone by the single-page application, which uses JavaScript and dynamic HTML, to call a servlet. What we need to consider for our remote evaluation system is (1) stable multi-user connection, (2) realization and operation of robot's choreography, (3) incorporation of questionnaires, (4) select the image format and parameters to realize sufficient resolution image and communication speed (fps), and (5) how to realize stable operation for a long time under the environment defined in (4). The followings are detailed explanations:

(1) For a stable multi-user connection, it is basic to reduce the amount of image communication on the network. Therefore, in order to reduce the access between the robot/server and avoid accesses by multiple evaluators as much as possible, we separate the manager page and the user page for evaluators, and allow the user pages to accesses the cache only.

(2) For the robot's choreography and its operation, multiple evaluators can observe the robot's movements at the same time, and only one evaluator can operate the robot for a fixed time. The robot autonomously plays a famous scene of about 30 seconds, as the same as Chapter 2. The operation of the robot is instructed by pressing a button in the direction of movement (up, down, right, left, stop). For example, by pressing "up" button, the destination feature point number which is appropriate to what "up" means is transmitted to the robot. The movement of the robot moves along the trajectory that interpolates the current position and the feature point of the destination. Figure 6 shows the relation between the button label and destination feature point in the actuator's movement of OSONO's choreography. When the robot receives the operation instruction, it feeds back by displaying a small mark on the image transmitted from the robot side. The robot's choreography is generated by interpolating the feature points, which are correspondent to characteristic scenes [2].

(3) For the incorporation of questionnaires, the questionnaires are implemented using Google forms with a reduced number of questions.

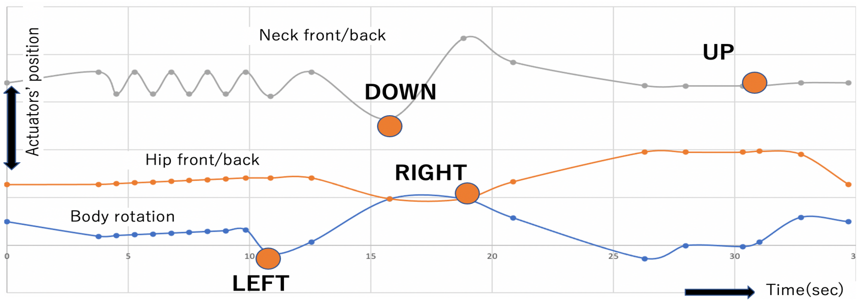

Figure 6: Button labels and destination feature points (red circle) in OSONO's choreography

(4) For the image communication speed, it is necessary to secure sufficient resolution image and frame rate for remote evaluation. Considering the latency, we choose JPEG compression and 640 x 240 (about 50 KB/frame), and expand the frame 2 times to a vertical direction at the browser. Increasing the image communication puts pressure on the network, finally, we choose 8 fps as the base frame rate, and we vary the rate depending on the network load. Currently, movie and/or real-time image communication is one of key technologies, such as Video Tag support in HTML5, public availability of H.264 implementation as OpenH264 [21] by Cisco Systems, and higher compression methods by major players. Since they have a higher compression rate than JPEG, we will consider to use in the next Chapter.

(5) For stable operation for a long time, it is key to grasp the outline by performance measurement and performance estimation, and investigate the cause in detail when a trouble occurs. To resolve this issue, we newly introduce a mechanism to monitor the communication status by measuring the disturbance of image data reception interval from the robot and accessing the image data from the PC. These let us adjust in real-time, and respond to wide load fluctuations for a few hours. This is very effective technic to realize robustness.

Since this architecture can manage multiple robots by name using RSNP, it can be realized by connecting multiple robots to a server and displaying multiple display/operation screens on a PC.

(6) Since it is common in both of Basic Remote Evaluation System and the extended Remote Evaluation System, we cover the delay in Chapter 6.

4.3 Stable Operation

In this remote evaluation system, the issue is how to increase the transfer rate, send high-quality images, and operate stably within the range that does not exceed the capacity and performance in order to ensure the sense of reality. This section considers the phenomena that hinder stable operation, their factors, and some solutions.

Trouble phenomena that are likely to occur are (1) slow image update during image transfer/display, (2) image transfer/display stops, but recovered when image transfer speed is slowdown, or when the display program is restarted. There are also, (3) the connection between the robot and the server is cut off, the robot server goes down, (4) the connection between the robot servers is cut off, but remain the image transmission continues, (5) when reconnecting, a series of old images is continued to display.

These are due to network and CPU overloads rather than specific library failures or communication method issues. However, the causes are related to various factors such as network fluctuations, PC hardware (CPU, USB, LAN adapter), OS, web server, browser, various libraries, application methods, and image compression. In addition, troubles occur 5 minutes to 1 hour after startup, and error messages often do not appear, resulting in poor reproducibility. Furthermore, in the system, there are many performance improvements, buffering and recovery mechanisms such as asynchronous operation, queues, buffering and retransmission of communication data, and there are many tuning parameters, making it difficult to understand the causal relationship.

To investigate the cause of these troubles, the following are effective: (1) Estimate network usage, (2) Change asynchronous operation to synchronous operation to measure processing time, (3) Visualize suppressed error messages (JavaScript, RSNP library, HTTP log), (4) Visualize hidden processing time such as Dynamic HTML in JavaScript). It is also effective to tune the parameters of the image transfer queue so that delays do not accumulate (an error occurs instead). As a result, it was found that the problem can be solved by changing the image transfer speed.

However, network and CPU loads are changing anytime and need to be monitored. Image transfer is implemented as the asynchronous operation, so it is difficult to measure the processing time. We found that the callback interval, called when the image is uploaded from the robot to the server, and the servlet call interval for the PC to acquire the image to the server shows the network and CPU overload status. As a result, we can monitor these values and realize stable operations.

5 Extended Remote Evaluation System

In order to clarify the effects of dynamic image quality and the effects of techniques with Ningyo Joruri, based on the basic evaluation system, we design and develop an extended remote evaluation system. Therefore, we improve the dynamic image quality, change button labels, add sounds of "shamisen (string)" and the narrations of "Tayu"[††] by using real-time audio transfer.

5.1 Improving Dynamic Image Quality

In this extended system, we try to adopt MPEG-4 AVC for smoother display using open source. At the same time, the communication traffic can be greatly reduced. In fact, in the case of Motion JPEG, assuming 50 KB per screen, and 8 fps as the transfer rate, if 10 persons access the system, the traffic will be 32 Mbps. On the other hand, in the case of MPEG-4 AVC, the Traffic will be 4 Mbps, even if the frame rate is 30 fps. Therefore, it is suitable for observing and evaluating the movement of robots. However, current open source-based tools take at least 1 second to acquire image files. This means that it is not suitable to operate OSONO's shorter choreography 10 to 15 seconds long.

[††] One narrator tells the story of Ningyo-Joruri. This includes describing the scene, to reciting the lines of each character.

The implementation is as follows: (1) The image acquired by the USB camera is compressed with MPEG-4 AVC of 30 fps at 1 second interval and the MP4 file is created on the robot side. Compression is realized by python3 OpenCV. Instead, it can be implemented by raspberry PI 3 with Java [13]. (2) Transfer the file to the server with RSNP's DistributeCameraImage, and cache on the server to accommodate access from multiple persons. (3) From the PC browser, access the server at 1 second intervals (or less to improve latency) by JavaScript, acquire video files, and play only newly generated files using the Video Tag function of HTML5. When displaying the image, z-index property in CSS can be used to fix two viewing areas to the same screen area, and a double buffer allows to play without the video is choppy. It is important to find that, as this implementation proves, it is possible to incorporate the high compression technology such as MPEG-4 into research systems by using open source.

5.2 Ningyo Joruri's Performance Effects

In order to verify the effectiveness of the performance with referring to "Ningyo Joruri," we added such performance elements to the control button labels, as audio, and feedback of the control buttons.

(1) Adding performance directing elements to the control button labels

Ningyo Joruri acting scripts are implicitly punctuated with seven-five syllables. For example, the beginning part of ”OSONO no KUDOKI”, a translation to English is “Hanshichi-san, where are you and what are you doing, now?”, can be punctuated as ”I ma go ro wa”, ” Han shi chi sa n”, “do ko ni do u shi te”, and “go za ro u zo” with five, five, seven and five syllables [14][15]. Label each of them as a button, and set it to changes current performance to the corresponding choreography and transition the performance when the button is pressed. At the same time, play the corresponding “shamisen”, narration, and the lines. Figure 7 shows the relation between button labels and actuator’s motion in the part mentioned above.

(2) Adding sounds of “shamisen” and the narrations by “Tayu”

In the real Ningyo-Joruri, a puppet performs with the narrations by “Tayu” and the sounds by “Shamisen”. To achieve this effect, the implementation is the following: (a) the sound data of the “Tayu” and the "shamisen" are played on the robot side when the corresponding choreography is started by operating the buttons. (b) This is routed to the robot's input processing by a virtual

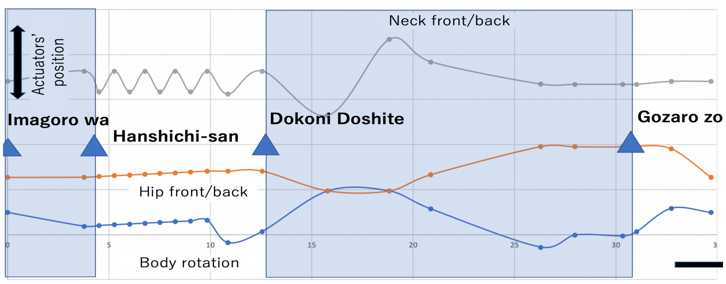

Figure 7: Relation between button labels and actuators motion
in the beginning of OSONO no KUDOKI

sound device. (c) The WAV file is created every 150 msec using python3's pyAudio. We used 16 bits PCM with a sampling rate of 8000 samples per second [16]. (d) This file is transferred to the server with RSNP's DistributeCameraImage protocol and cached for access by multiple audience. (e) From the audience's PC browser, JavaScript is used to access the server through a servlet at 75 ms intervals to retrieve WAV files, and only newly created files are played when they are ready to be played. The double buffer method is used for playback, so that the breaks are not so noticeable.

(3) Feedback on button pressing

It takes about one second from the time the button is pressed until the robot actually starts moving and reaches the designated position. To prevent this delay from confusing the evaluator, in spite of the Basic Remote Evaluation System already feeds back the button operation to the robot's display screen when it is received, we explore more effective methods by (a) making it a larger display, (b) lighting the lanterns and (c) playing the shamisen sound.

6 Demonstration Experiment

Using the two remote evaluation systems, the basic version and the extended version, described in Chapters 4 and 5, we verify in 6.1 whether the remote evaluation is as good as the local evaluation. In addition to the overall evaluation in 6.2, the effects of dynamic image quality and production elements are clarified. Furthermore, we propose guidelines for effective system design for online exhibits in 6.3, based on the previous sessions.

6.1 Remote Robot Evaluation

We verified the prototype remote evaluation system and conducted a demonstration experiment to evaluate the choreography using it. The demonstration experiment was conducted during one hour demonstration time of the actual online event, "Bay Area Robotics Forum 2021" held on 2021/1/27 from 14:00 to 18:00. OSONO robot, with the same setting as test case B in Table 1, was evaluated with the event participants as the evaluators. This confirmed that the remote evaluation system met the original design concept (Figure 8).

We deployed Apache web server and Tomcat on the server (Xeon Processor 12 cores 24 threads), and we installed a robot (Mac book pro) at Aoyama Gakuin University, and cooperated with the event system portal at Shibaura Institute of Technology. Video distribution was set at 640 x 240, 8 fps. The communication status was monitored by the load monitor mentioned in 4.4, and HTTP request was monitored using Apachetop. As a result, continuous operation was performed from 15:00 to 18:00, the system was stable, and there were no robots down nor server down. The network speed is 970 Mbps for uplink and 650 Mbps for downlink.

The questions in the questionnaire are divided into those for general evaluators who do not operate the robot and those for evaluators who operate the robot. The former is almost the same as the questions for 2.3 and 3.1.

Actual questions are： (1-1) OSONO is attractive as a robot, (1-2) OSONO is familiar, (2-1) No discomfort in the choreography, (2-2) OSONO’s choreography is attractive. The evaluation

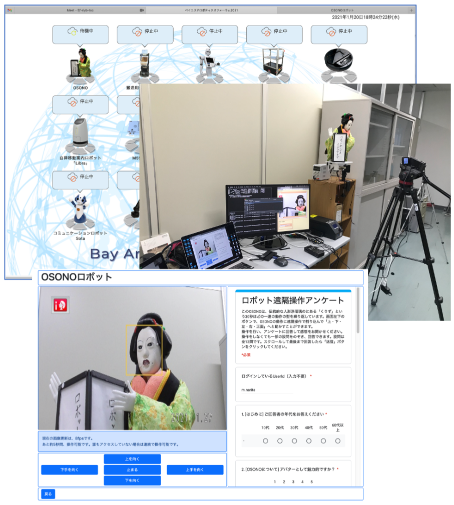

Figure 8: The remote evaluation system during demonstration experiment

points are 1 (affirmative) to 5 (negative). There are 11 evaluators (university officials and business people, 40-60 years old) during the demonstration time. The results are shown in table 4.

The questions for the evaluator who operated the robot are: (3-1) Do you feel like you are manipulating a puppet by yourself? (3-2) Is the choreography of the robot easy to understand? (3-3) Is the operation screen easy to understand? (3-4) Did you enjoy the operation? There were 6 respondents (54%). The results are shown in table 5. What can be seen from the verification results is (a) very similar to the evaluation result by actual robot（ordinary person） in Table 2 of Chapter 3, and (b) different from the evaluation result by actual robot (Puppeteers, Expert) in Table 1 of Chapter 3. In other words, since the remote evaluation system shows the same tendency as the actual robot, it can be said that this remote evaluation system can be sufficiently used for the evaluation of choreography, even the movement of the 3D robot is represented on the 2D display.

Table 4: Evaluation result by Basic Remote Evaluation System

Question	Score
(1-1) OSONO robot is attractive	2.1
(1-2) OSONO is familiar	2.6
(2-1) No discomfort in the choreography	2.5
(2-3) Choreography is fresh?	2.4

Table 5: Evaluation result of Operate robots through Basic Remote Evaluation System

Question	Score
(3-1) Do you feel like you are manipulating a puppet by yourself?	2.5
(3-2) Is the choreography of the robot easy to understand	3.0
(3-3) Is the operation screen easy to understand?	2.0
(3-4) Did you enjoy the operation?	2.2

Additionally, the result on "OSONO is familiar" might be improved if we have used the head model showed in Figure 4. (2) Reality seems to consist of resolution (640 x 240) and transmission speed. Since there is no comment on the former in the questionnaire result, it will be true that there is no big problem in the image resolution. In the latter case, there is a comment on the image update interval, and it seems that the reality can be further improved by improving the frame rate. (3) Regarding the operation of the robot, the evaluators seemed to be able to realize the feeling of moving it by themselves and enjoyed the acting remotely. However, the evaluation value for robot movement is 3.0 (median). This may be due to the delay described in 3.3, but it is necessary to consider ways to make it easier to understand the actions of the operator.

6.2 Evaluation on Dynamic Image Quality and Theatricality Performance Effects

(1) Experimental methods and Experiment systems

We conducted two experiments with the extended remote evaluation system, with ten undergraduate and master's students (in their 20s) in the Department of Mechanical Engineering, majoring in intelligent robotics, were selected as evaluators. One experiment is (a) Dynamic image quality described in 5.1. Another experiment is (b) Robot remote manipulation experiments with the performance effects described in 5.2. Since the evaluators had no knowledge of Ningyo Joruri, a 10-minute explanation of Ningyo Joruri was given before the experiment.

We deployed our server applications on Apache web server and Tomcat on the server (Xeon CPU 12 cores 24 threads) at Aoyama Gakuin University. Evaluators' PC and a robot side application on Mac book pro, at Shibaura Institute of Technology, are connected with WiFi via the Internet (SINET) to Aoyama Gakuin University's server. As dynamic image transfer, for experiment (a), MPEG-4 AVC (640 x 480 pixels 30 fps) is adopted, and for experiment (b) Motion JPEG (640 x 240 pixels 8 fps) is adopted. Experiments took less than one hour in total. Ten PCs are connected to the system at the same PC with no problem.

Table 6: The delay at each event

Event	Local(msec)	WAN(msec)
From pressing a button to Image feedback (JPG)	237	345
From Pressing a button to recognizing lighting the lantern at a user PC	321	347
From Pressing a button to recognizing the audio playing at a user PC	248	375
From Pressing a button to recognizing an actuator's starting at a user PC	380	552
Image transfer in Motion Jpg (8 fps)	190	210
Image transfer in Mpeg-4 AVC (30 fps)	1390	1576

The delay at each event is shown in Table 6. The delay caused by Motion JPEG（it is the same as the basic remote system case, as the extended remote system case）, the delay from pressing the button to the image feedback, lighting the lantern, and starting the sound are all within 0.4 sec even in WAN environment. Also, actuators start in about 0.55 second. Compared to 0.4 seconds [17], which is the standard latency for cell phones, these feedback delays are numerically acceptable, and the delay before the actuator starts working is not significant. Some fluctuation at the WAN environment is also considered to be acceptable. These values are an assumption for interpreting the results of the questionnaire.

On the other hand, the delay with MPEG-4 is about 1.5 seconds, and with the addition of the operation response, it will be around less than 2 seconds. Since, OSONO's choreography is 35 seconds, and it is divided into four sections, so one section is about 7.5 seconds. If the choreography about 7.5 seconds long is operated with a delay of around less than two seconds, it is considered to lose its sense of control. Therefore, we judged that OSONO operation experiments using MPEG-4 as a monitor would not yield good results, and we did not include it in our experiments.

(2) Evaluation of real-time display using MPEG-4 AVC

In this experiment, evaluators observe OSONO repeating the choreography in real-time using MPEG-4 AVC images, and answer a questionnaire at the same time. The questionnaire items were (1-1), (1-2), and (2-1) used in 6.1 for the general appeal of robots and pretending, plus (2-3) "Choreography is fresh? and a request for comments on the real-time video in a free-entry format. The evaluation points are 1 (affirmative) to 5 (negative). The results on (1-1), (1-2), (2-1), (2-3) are shown in table 7. On (3), we received two general comments on the video, one is that mpeg-4 video is smoother than Motion JPEG, another is that PC video processing is heavy.

Table 7: Evaluation result by Remote Evaluation System (Mpeg-4)

Question	Score*
(1-1) OSONO robot is attractive	2.2
(1-2) OSONO is familiar	3.3
(2-1) No discomfort in the choreography	2.8
(2-3) Choreography is fresh?	1.5

*A lower score means a better evaluation.

The answer to (1-1), (1-2), and (2-1) in Table 7 are almost equal to the remote evaluation of 6.1 (Table 4). This concludes that the smoothness of the MPEG-4 AVC video does not directly affect the robot's evaluation. On the other hand, 2-3 (freshness) is 1.5, which is higher than 2.0 and 2.5 in the case of Motion JPEG described below in 6.1 and 6.2. This indicates that using MPEG-4 at 30 fps causes the choreography looks fresh. Therefore, it is reasonable to use MPEG-4 for video display unless the delay is not essential.

(2) Evaluation on Ningyo Joruri's performance effects

Using the Motion JPEG, OSONO repeating the choreography is displayed in real-time, and the evaluator operates the robot, then answers the questionnaire. The questionnaire items are (1-1), (1-2), (2-1) and (2-3), which are related to the general attractiveness of the robot used in 6.1 and 6.2(1), plus (3-1), (3-2), (3-3), (3-4) in 6.1 which related to the evaluation of the manipulation, and added what is related to the performance effects and the manipulation in detail. Actually, the following items are added: (4-1) Were you aware of the screen marks as feedback for operation? (4-2) Did you pay attention to the shamisen and voice (the lines)? (4-3) Did you notice the delay in the lighting of the lanterns? (4-4) Were you bothered by the delay of the screen mark? (4-5) Which of the feedbacks did you pay particular attention to? (5-1) Did you operate with the goal of changing the lines? (5-2) Was it appropriate to use the change the lines as the goal of the operation? (Only for persons who answered yes in 5-1) (5-3) Did you operate with the goal of changing the choreography? (5-4) Was it appropriate to use the change the choreography as the goal of the operation? (Only for persons who answered yes in 5-3) The evaluation points are 1 (affirmative) to 5 (negative). The results correspondent to table 7 are shown in table 8. Additional Performance effects Evaluation result is shown in table 9.

Table 8: General Evaluation result by Extended Remote Evaluation System (Operation)

Question	Score*
(1-1) OSONO robot is attractive	2.1
(1-2) OSONO is familiar	2.9
(2-1) No discomfort in the choreography	2.6
(2-3) Choreography is Attractive and fresh?	2

Table 9: Performance effects Evaluation result by Extended Remote Evaluation System

Question	Score*
(3-1) Do you feel like you are manipulating a puppet yourself?	2.5
(3-2) Is the choreography of the robot easy to understand?	3
(3-3) Is the operation screen easy to understand?	1.7
(3-4) Did you enjoy the operation?	2.0
(4-1) Were you aware of the screen marks as feedback for operation?	52%
(4-2) Did you pay attention to the shamisen and voice (the lines)?	76%
(4-3) Did you notice the delay in the lighting of the lantern?	40%
(4-4) Were you bothered by the delay of the screen mark?	30%
(4-5) Which of the feedbacks did you pay particular attention to?	Screen mark 70% Shamisen/the lines 20% Choreography 10%
(5-1) Did you operate with the goal of changing the lines?	50%
(5-2) Did you operate with the goal of changing the choreography? (Only for persons who answered yes in 5-1)	1.4
(5-3) Did you operate with the goal of changing the choreography?	60%
(5-4) Was it appropriate to use the change the choreography as the goal of the operation? (Only for persons who answered yes in 5-3)	1.7

- General

The answer to (1-1), (1-2), and (2-1) in Table 8 are almost the same as the general evaluation in Table 4 and 6.1, indicating that robot OSONO is evaluated as attractive. On the other hand, when we look at Table 8 in details, the same evaluator improved 0.1 to 0.4 compared to Table 7. Table 8 shows the results of the questionnaire after the respondents had operated the robot after the answers in Table 7. It can be said that the longer time spent in the robot operation has a significant impact on the improvement of overall evaluation, including attractiveness.

The response to the audio was high, with 76% of the evaluators being aware of "Shamisen" and voice (the lines). However, 70% of the evaluators confirmed the feedback of button operation with screen marks, not voice, thus screen marks are suitable for this purpose. Evaluators do not feel the delay on lighting the lanterns and the screen feedback. However, evaluators are not very aware of lighting of the lanterns.

- Robot operation

The answer to question (5-1) was 2 points, the same as in 6.1, and we can understand that the evaluators feel that they are operating the robot by themself. The answer to (5-2) is the same as in 6.1, this indicates that the operation instructions by the lines and phrases are as easy to understand as instructions by the direction of movement. There is room for improvement. For example, in 6.1, the choreography corresponding to "down" and "right" instructions might be too close in time (in Figure 6). In 6.2, it might be difficult to find the starting point in choreography. Actually, in Figure 8, although "Imagoro wa" is the beginning, the corresponding choreography is more monotonous than the ending part. Making the beginning phrase faster and the end phrase slower, the choreography will be clear. The choreography "Kuriuzu" in the third phrase is both a very famous and complex form, and it would be more attractive if the form were extended to the 4th phrase for greater visibility.

The score for the answer to question (5-3) was 1.7, which is an improvement in the clarity of the operation screen compared to 2.0 of the answer in 6.1 (Table 5). This is caused by various types of feedback such as screen feedback. The answer to the question (5-4) was 2.0, which is an increase in enjoyment of the operation compared to 2.2 of answer in 6.1 (Table 5). This may be due to the contribution of the entertainment elements of shamisen and the narration and the lines.

- Operation Goal

Question (5) asks what the operation goal was set for the completion of the operation. 90% of the evaluators chose the goal of having the lines or choreography change with the operation, and were satisfied (1.4 to 1.7) enough with their choice of the goal. 50% of the evaluators targeted the changes in the lines as the goal of the operation, and as a result, they answered their choice of goal was appropriate (1.4). On the other hand, 60% of them targeted changes in the choreography, and as a result, they answered their choice of target was appropriate (1.7). Therefore, for evaluators who target the lines, improving the sound quality and choppy, and improving the clarity of the correspondence between sound and choreography will improve operation feel for the puppet better. And For evaluators who target the choreography, improving the clarity of the choreography will lead to a better operation feel for the puppet.

6.3 Summary and Guidelines for Effective System Design in Online Exhibitions

The remote evaluation system evaluated that OSONO was attractive and that the sense of control of the puppet could be fully obtained. In this section, as an exhibition for cultural experience by combining traditional culture and information technology, we discuss the viewpoint of extended media and online exhibition.

There are [18] and [19] related studies that convey the values of culture through the exhibitions that combine augmented reality media with the things handled by the specialists and traditional cultures in a specific region such as intangible cultural properties. In [18], the authors created a system in which the image projected on the furoshiki changes due to the act of wrapping the object in the furoshiki. Thereby, it was confirmed that even people who are not familiar with furoshiki can feel the values of behavior and things by experiencing a traditional culture with interest. In [19], in order to convey the culture of worship at a shrine, the authors created a work that uses video projection based on body movements and object recognition in a fusion with a picture scroll. In addition, there is a study [20] that introduces physical interaction into the museum experience.

The remote evaluation system described in this paper not only conducts evaluations, but also has provided a simulated experience of a puppeteer that only an expert in Joruri puppets could handle by combing traditional culture with robotics and Internet distribution at the same time. We summarize the results of 6.1 and 6.2 and propose guidelines for effective system design in online exhibitions.

- OSONO is rated as attractive overall, regardless of whether it is evaluated by the general public, by puppeteers. The remote evaluation is able to obtain the similar results with the face-to-face evaluation through our remote evaluation system. The results do not change whether the video transfer rate is 30 fps or 8 fps. On the other hand, the amount of time spent touching and operating OSONO has a significant impact on the overall evaluation, including attractiveness. Explaining in advance to an audience unfamiliar with Ningyo Joruri and increasing the time spent touching the puppets will result in a higher evaluation.
- When viewing, using high compression technology such as MPEG-4 for high frame rate video transfer will give you a fresh feeling. Therefore, Motion JPEG, which has small delay but requires wider band width in the network communication, should be used only when audience manipulate the robot.
- Audiences are sensitive to audio, such as "shamisen" and voices (the lines or narration). The use of "shamisen" or voice (dialogue or narration) in combination with the audio makes the display and manipulation more enjoyable, and helps to increase the amount of time the audience can interact with the robot. None of the feedback delays are a problem, but the screen feedback is highly appreciated.
- The sense of control of the puppet can be fully obtained through acting instructions by buttons. When the audience manipulates the robot, they use the changes in the lines and the choreography of the robot as guides. Instructions to the robot's choreography can represent by pointing the direction of movement, but could be difficult to distinguish sometimes. Thus, instructions by the lines/phrases are more appropriate. However, in order to make it easier to understand which gesture corresponds to which line, gestures should be clarified and simplified to the extent that the quality of the gesture is not compromised.

7 Conclusion

In this paper, we overviewed OSONO project, a robot based on Joruri puppets, and compared the results of evaluations by a group of puppeteers (experts) with existing evaluations of ordinary people. As a result, it was verified that OSONO is "attractive" to a wide range of people, including not only the general public but also puppeteers. At the same time, we verified that the quality of the robot and its choreography could be realized even with a small number of actuators in the implementation of the choreography. On the feed backs from experts at the verification, we enhanced the quality of the robot and choreography using this information.

In addition, we tried to evaluate the robot using a prototyped remote evaluation system including robot operation with the open software, and verified that the same evaluation results as face-to-face evaluation can be obtained even with remote evaluation including acting operations. In addition, we analyzed and evaluated the impact of the real-time video, and the effectness of the directing elements in the performance with referring to Ningyo Joruri such as button labels, shamisen, narration and the lines. Finally, we proposed a guideline for designing an effective system for an online exhibit that would bring out the best of the robot with referring to Joruri puppet.

In the future, we try to develop an online exhibition system that emphasizes video quality and acting elements, and to realize more attractive service robots.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 20K12011.

References

[1] Masahiko Narita, Sachiko Nakagawa, Development of OSONO, a service robot with reference to “Joruri puppet”, and its Choreography, IJSKM International Journal of Service and Knowledge Management, pp.54-70, 2020

[2] Masahiko Narita, Sachiko Nakagawa, Physical properties of service robots which is referring to Joruri puppets, The IEEE/IIAI International Congress on Applied Information Technology (IEEE/IIAI AIT 2019) pp. 158-162, 2019

[3] Toshikazu Kanaoka, et al., Designing a Motivational Agent for Behavior Change in Physical Activity, CHI EA '15: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2015), pp.1445-1450, 2015.

[4] Masahiro Shiomi, et al., “Effects of an Appearance and a Size for a Robot Signage," Forum on Information Technology proceedings (FIT2010), 9(3), pp.83-88, 2010 (in Japanese)

[5] Chao Shi, et al, Field Trial for Social Robots that Invite Visitors to Stores, Journal of the Robotics Society of Japan, Vol. 35 No. 4, pp.334~345, 2017

[6] Zhe Cao, et al., “Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields”, 2017 IEEE Conference on CVPR, pp.7291-7299, 2017.

[7] Council for Promotion of Universal Future Society, “Robot showcase,” 2021, url https://www.youtube.com/playlist?list=PLGpGsGZ3lmbD_5TJz-bsdyFE1krx2Plc3, online

[8] Masahiko Narita, et al., Demonstration experiments of a distributed questionnaire service using multiple robots with the aim of marketing at a tourist site, International Journal of Smart Computing and Artificial Intelligence (IJSCAI),2018, Vol.2, No.1, pp.1 - 21, 2018

[9] Masahiko Narita, Sachiko Nakagawa and Toru Izui, “Evaluations of service robots referring

to Joruri puppets, and a remote evaluations system", Proceedings of 12th International Conference on E-Service and Knowledge Management, 2021

[10] Masahiko Narita, et al., Development Framework for Non-experts targeting the Network Robot Service, Journal of the Robotics Society of Japan, Vol.33, No.10, pp. 807-817,2015/5

[11] RSi-Robot Service initiative: http://robotservices.org/ [Online]

[12] M. Narita, Y. Murakawa, C. Akiguchi, Y. Kato, T. Yamaguchi, "Push communication for network robot services and RSi/RTM interoperability," FUZZ-IEEE 2009. IEEE International Conference, Korea, August 20-24, 2009. pp. 1480-1485, 2009.

[13] Satoshi Okano, Koichiro Kato, Yukihiro Nakamura, Nobuto Matsuhira and Masahiko Narita, "Development and demonstration of a general-purpose communication unit for a robot cooperation network," Journal of the Robotics Society of Japan, Vol. 39, No. 10, pp. 973-980, 2021

[14] Masahiko Narita, "Extraction of Choreography for Service robots from a Ningyo Joruri's Acting script by the Associative model," Proceedings of IIAI AAI 2021-Winter (11th International Congress on Advanced Applied Informatics)

[15] Masahiko Narita, Sachiko Nakagawa "Systematizing the choreography of Ningyo Joruri by an Association Model from the perspective of a Service Robot" Journal of the Robotics Society of Japan, 2022

[16] Hideya Osawa, Kensuke Asakura, Noriko Obara, Naohiro Fujita, Ken Sato, Sachiko Nakagawa and Masahiko Narita, "Robot service platform for integration of voice communication with robot control by RSNP extended," Journal of the Robotics Society of Japan, Vol. 33, No. 2, pp. 103-114, 2015

[17] ITU-T Recommendation G.114: One-way transmission time, 2003.

[18] Kei Kobayashi, Kazuma Nagata, Soh Masuko, Junichi Hoshino. FUROSHIKI: Augmented Reality Media That Conveys Japanese Traditional Culture. ACM SIGGRAPH VRCAI 2019, pp1-5, 2019.

[19] Kei Kobayashi, Junichi Hoshino. Augmented Reality Media for Cultural Experience in Shrines. IFIP International Conference on Entertainment Computing (ICEC 2017), pp375-378, 2017.

[20] Satoru Tokuhisa, et al. A Service Design Practice using Special Exhibition "Map of Japan created by Ino Tadataka" at Tokyo National Museum and Museum Theater. Transactions of the Virtual Reality Society of Japan (TVRSJ) Vol.20 No.1 pp.3-14, 2015. (in Japanese)

[21] Open H264 https://www.openh264.org, online