Journal of Information Processing Vol.25 352–360 (Jan. 2017)

[DOI: 10.2197/ipsjjip.25.352]

Regular Paper

Embodiment of Guidance Robot Encourages Conversation among Visitors

Kohei Matsumura[1,a)] Yasuyuki Sumi[2,b)] Takumi Gompei[2]

Received: August 5, 2016, Accepted: February 9, 2017

Abstract: In this paper, we propose a novel way to encourage visitors to share their experiences and interests in exhibition spaces. Visitors may have experiences in an exhibition and become aware of meanings of the exhibits and/or relationship among them. We believe that sharing the experiences of visitors will enhance the exhibition experience for subsequent visitors because shared experiences may include fascinating topics. To acquire experiences of the visitors, we used "PhotoChat," which is an in-house photo communication software. PhotoChat is capable of communicating with others by taking photographs and adding annotations to each photograph. It also records the locations for coordination between the photographs and the statistical information contained in the annotations. Since PhotoChat is designed for realtime communication, in this study, we introduce a robot that inhabits the exhibition space. The robot is always on the PhotoChat and acquires all data on PhotoChat. The robot, thus, is capable to know what a visitor communicate with others on PhotoChat and to share them with subsequent visitors. The robot can also use bodily actions to express instructions to the visitors. We developed a system that integrates PhotoChat into a robot. We also implemented robot behavior (i.e., bodily actions and motions) that includes recommendations for photographs taken by others. That is, the robot communicates with human using both PhotoChat and its body. We held workshops to perform data collection and manually classified the data into three content categories. We then performed experiments using the developed system to distribute the classified content. The results showed that the robot's physical behaviors encouraged conversations between the visitors based on provided topics.

Keywords: guidance robot, content distribution, embodiment, exhibition, conversation

1. Introduction

This paper describes our work towards a realization of an exhibition guide robot that helps visitors to share their experiences.

In exhibition spaces such as a museum, visitors have experiences with exhibits and become aware of meaning of the exhibit and relationship among them. They also take pictures, notes and even sketch the exhibits to memorize their experience. They ocassionaly share these reminders with their family/friend and may talk about their experiences. Althourgh these seem to enrich the experiences of subsequent visitors, it might not be shared with the visitors.

We developed "PhotoChat" that supports casual conversation among the users by allowing them to take and share photographs and add hand-written notes to the photographs [1]. PhotoChat enables visitors of an exhibition to share their personal experience with accompaning persons in realtime. However, PhotoChat is designed for realtime communication. That is, it is difficult to share experiences with subsequent visitors even if these enrich the experience of subsequent visitors.

In this study, we tackle mainly two problems on sharing experiences in an exhibition, one is that difficulty with sharing past PhotoChat conversation and the other is that unawareness of shared contents. For the fitst problem, as we described above, it is difficult to share the past conversation on PhotoChat because PhotoChat is designed for realtime communication and the subsequent visitor cannot grasp the context of the conversation. To understand the experiences of past visitors, knowing its context seems to be important. The context may include where/when a photo was taken and who took a photo. For the second problem, if shared contents (i.e., experiences someone had in past times) are not related to nor colocated with the exhibit a visitor currently watches, it should be unaware/ignored because the visitor cannot grasp the relationship between the exhibit and shared contents.

To overcome above two problems, we introduce a robot that inhabits the exhibition space. The robot is always on the PhotoChat and acquires all data on PhotoChat.

The robot, thus, is capable to know what a visitor communicate with others on PhotoChat and to share them with subseuent visitors. PhotoChat handles not only photos with comment but also its context including time and location coordinates respective exhibition area. The robot can also use bodily actions to express instructions to the visitors. We expect that bodily actions of the robot persuade visitors to aware shared contents. The actions also help the visitors understanding of the context of shared contents.

Figure 1 shows a typical scene where our system is working in a small exhibition space. Individual visitors can freely enjoy their tour of the exhibition while being connected with PhotoChat. The robot located at the center of the exhibition space observed

1 Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
2 Future University Hakodate, Hakodate, Hokkaido 041–8655, Japan
a) matsumur@acm.org
b) sumi@acm.org

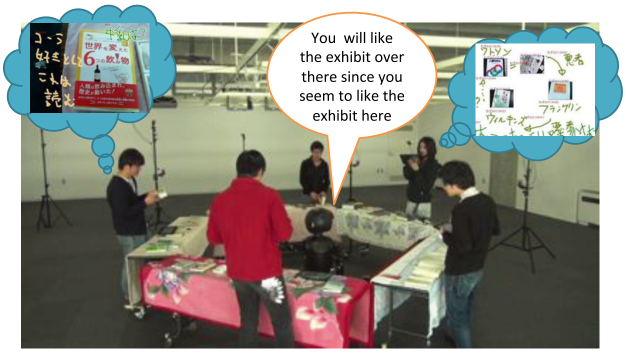


Fig. 1 Robot guidance based on the PhotoChat conversations among the visitors.

the individual visitor's touring histories and the PhotoChat-based conversations among them. The robot is then able to gather an overview of the exhibits, e.g., it can detect popular topics with regard to certain exhibits and observe relationships among the exhibits from the accumulated data. The robot can then provide new visitors with exhibition guidance based on the previous PhotoChat conversations.

In this paper, we investigated the effects of the robot's embodiment, i.e., its speech and gestures, on the visitors' acceptance of the robot and the interactions among the visitors. Contributions of the paper are as follows:

- We built a system that is capable to share visitors' experiences with subsequent visitors in an exhibition.
- We found that our system successfully affects the visitors' experiences at an exhibition.
- We also found that our system may persuade conversation not only on PhotoChat but also in real world.

2. Robot Guidance Based on PhotoChat Conversation among Visitors

The proposed robot guidance system supplies users with context-aware information provided by a robot that is networked with both environmental sensors and a hand-held systems that are used by the individual users.

We chose exhibition spaces to be our experimental field. The individual visitors carry hand-held systems (running PhotoChat) that enable them to share their interests and new findings instantly during exhibition tours with other visitors. Their touring histories (i.e., their current position and the time spent at each exhibit) are recorded using a location system, i.e., an optical motion capture system.

At the center of the exhibition space, we provide a guide robot, which acts on behalf of the human facilitators, and speaks to the visitors to suggest topics related to the individual exhibits and to recommend other exhibits to the visitors. The guidance timing and addressees are calculated based on the touring histories of the individual visitors. The robot sends a photograph to all PhotoChat terminals to serve as guidance content at times when it is speaking to a particular visitor. The PhotoChat photo used has been selected from stored photos taken by previous visitors; e.g., a photograph showing findings on and conversation with regard to certain exhibits by previous visitors is conveyed, based on the similarity of their contexts.

One important characteristic of our system is that we use knowledge from the experiences of the visitors (the system users), and not knowledge that was prepared by system designers beforehand. Another important characteristic is that our system embeds the guidance content within the physical context using a robot embodiment, and not simply via personal mobile systems. Our research interests also include the social effects of the guidance provided (i.e., acceptance of the guidance provided and initiation of interactions among the visitors) by the robot embodiment (including the effect of its approach, speech, head gestures, and hand gestures).

3. Related Work

Building robots that can move autonomously based on an understanding of their environment and that can then assist people has been a long-standing goal for researchers in artificial intelligence and robotics. Many works on guide robots have been conducted at venues such as museums [2] and offices [3]. This pioneering research into guide robots focused mainly on autonomous movement based on a visual understanding of the surrounding environment and spoken dialogue with people. Also, Tezuka et al. and Nakamura et al. addressed a framework for network robots that provide services to people [4], [5].

Recently, many works have focused their research in another direction, i.e., on social interactions between people and robots. For example, Kanda et al. [6], [7] presented a robot that could act as a social partner for children, that had stayed for long periods and had communicated with children at a primary school and a science museum. The research group of Yamazaki and Kuzuoka [8], [9], [10] analyzed the effects of the nonverbal behavior of museum guide robots such as stance, the coordination of head movements with speech, and their paralanguage based on a sociological interaction analysis.

The robot guidance parts of these works was based on guidance content that was prepared beforehand by the researchers and the museum curators. In contrast, our system attempts to use the visitors' collaborative experiences and findings during their tours to realize a robot that can offer a medium to help the visitors to share their knowledge and findings.

One noteworthy attempt to collect real-world news by talking to people nearby was proposed by Harada et al. [11] to act as a robot journalist. The robot automatically recognizes signboard and poster like objects, as they assumed these contain news, and makes a report using image processing technologies. However, it is still difficult to determine significant scenes from visual processing alone. We use PhotoChat [1] to ease the exchange of personal interests and findings among the exhibition visitors. We expect our approach to raise an intuitive exchange of knowledge that is embedded in the exhibition context.

Riether et al. revealed that human/robotic presence effects on human performance on both cognitive and motoric tasks [12]. This may implicate that mere robotic presence effects on visitors' experiences in museums. We previously proposed a method to encourage visitors to share their experiences by recording the social interactions among them and then virtually attaching first-person view videos and speech to the context in the exhibition

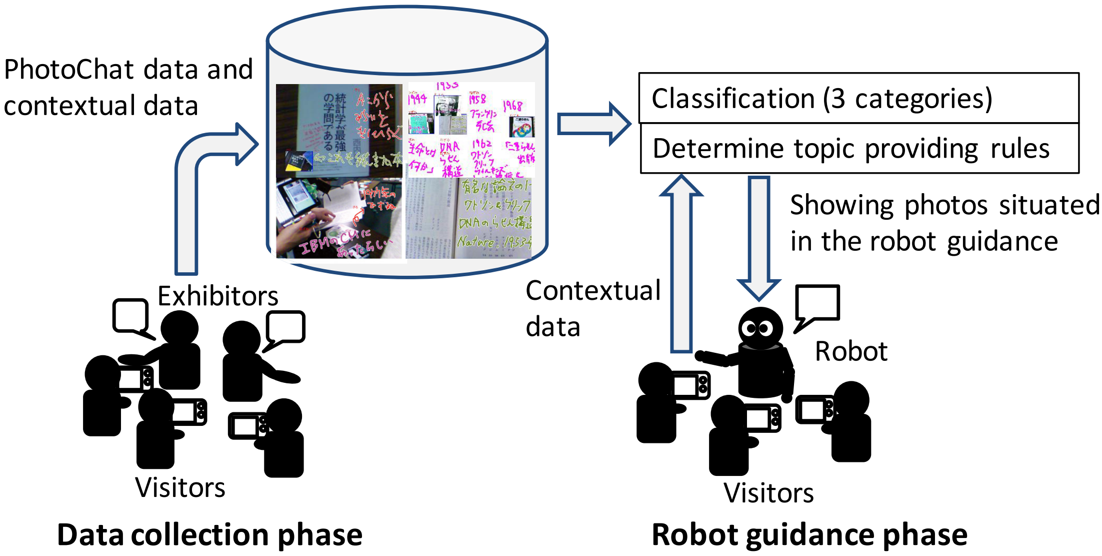


Fig. 2 Framework used for this study. The study consists of two phases. The first is the data collection phase, and the second is the robot guidance phase. The system collects PhotoChat conversation along with its contextual data. The system then classifies them into three categories. The robot will provide contents to visitors according to their behavior.

space [13]. In this paper, we use PhotoChat to enhance the visitors' casual record-taking and allow them to exchange their interests and findings by taking photographs and making hand-written notes on them. This paper investigates the effects of the robot's embodiment, e.g., its positioning, head gestures, and hand gestures, on stimulation of social interactions among the robot and the visitors [14].

4. Implementation

Figure 2 shows the framework used for this study. The framework consists of two phases. The first is the data collection phase and the second is the robot guidance phase. We held an exhibition of books as a space for the data collection phase, where we collected the "conversation" between the visitors. The participants held "conversation" via the PhotoChat system. This means the participants send messages that consist of photograph and/or drawing through the PhotoChat system each other as a conversation. The detailed descriptions of the system are described below. In the robot guidance phase, the robot distributed selections from the collected conversations as content for the guidance phase. We classified the collected chats on the basis of statistical data, such as the number of text annotations and/or the structure of the conversation.

4.1 PhotoChat with Location Aware Feature

As a first step towards system implementation, we held two workshops and collected information on the participants' interests and some insights into the exhibits. Figure 7 shows some snapshots taken of these workshops. There are five tables in the space and about ten books are on each table. The visitors freely picked up the books and had conversations about their interests and insights via PhotoChat.

Figure 3 shows an annotated screenshot taken from PhotoChat. PhotoChat is software that runs on mobile PCs that have a camera module and a pen input. Users can take photographs as they

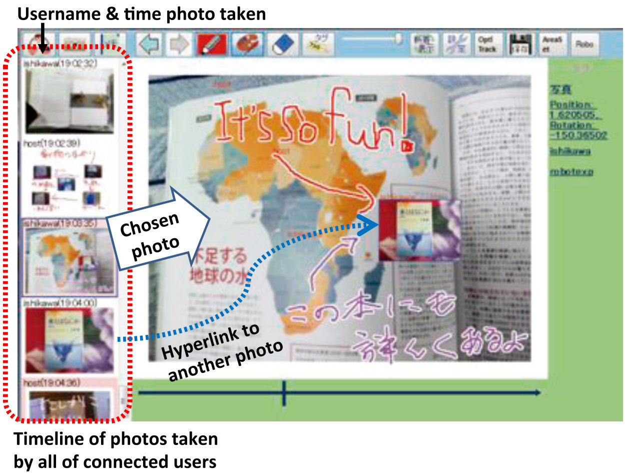


Fig. 3 An annotated screenshot of PhotoChat.

would with conventional digital cameras and write and draw on them via the pen interface. The photographs and the pen strokes are distributed to the other machines that are running PhotoChat by an ad-hoc wireless network in real time, so that the users can share photographs and conversations about them.

PhotoChat handles three types of data: photographs, handwritten strokes, and/or hyperlinks to other data. The typical usage is that a user firstly takes a photograph; secondly the user annotates the photograph by drawing; also, the user makes hyperlinks among related data into the photograph. As shown in Fig. 3, most of the PhotoChat display is laid out to show the camera viewfinder and the photograph browsing/notes. Thumbnails of all photographs taken by the connected users are listed on the left hand of the display, along with the times they were taken. Each thumbnail shows the user name and the time for each photograph. PhotoChat also has a unique function for hyperlinking of the photographs. The users can hyperlink between photographs by dragging thumbnails of the photographs and dropping them onto another photograph or a blank sheet. The users can then

Journal of Information Processing Vol.25 352–360 (Jan. 2017)

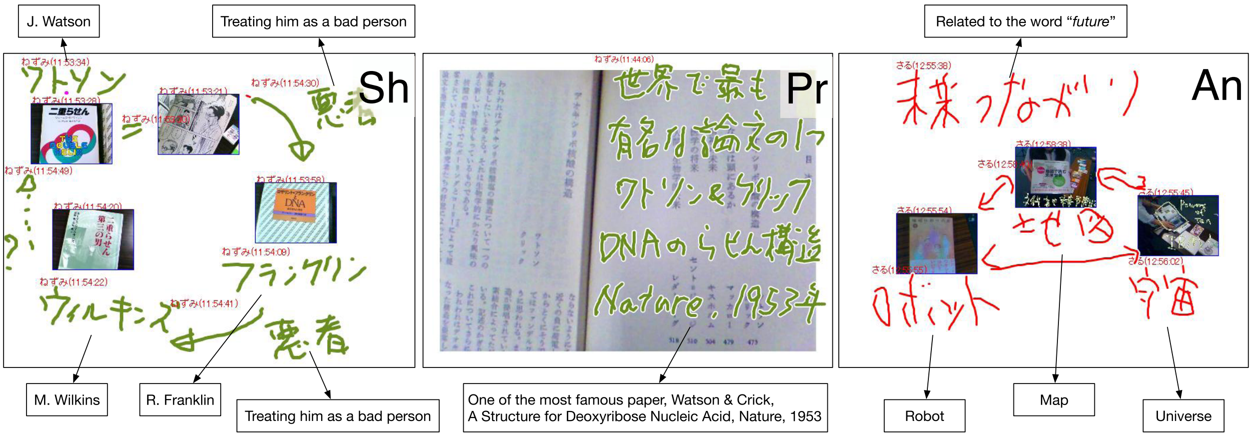


Fig. 4 Examples of the three categorized photo. Hand annotatations are written in Japanese. English translations (i.e., enclosed text) follow black arrows. **Sh** (left) indicates an example of the category "short description." **Pr** (center) indicates that of "providing topic." **An** (right) indicates that of "relationship between exhibits."

jump to the linked photographs by clicking on the thumbnails.

We collected the PhotoChat data, including the photographs that were taken in the workshop and the handwritten annotations on the photographs. In addition to these data, we also collected the metadata (i.e., the context) for each photograph, such as user name, location and time taken. We can estimate the areas that got the most attention, the stimulated conversations, and typical structures for the conversations from the context data of the conversations on PhotoChat. To gather the locations (i.e., areas) of the participants, we used a motion capture system (Natural Point Inc., OptiTrack).

The motion capture system recognizes the position of the participants with millimeter precision, but we uses the data to recognize the area of the participants. More specifically, we defined five areas according to each of five exhibitions, and the system recognizes whether the user is in the defined area or not.

4.2 Classification of Photographs

In this paper, we classified the annotated photographs on PhotoChat into three categories.

The first is the type where the photograph includes a "short description (**Sh**)" of the exhibit, e.g., a visitor abstracts the theme of the exhibit and annotates it on the photograph. **Figure 4** (left) shows the typical example of the category (**Sh**). That may include several photos of the cover of the book and may describe relationships among them with both text annotations and symbols.

The next category is the type where the photograph provides a detailed description of and/or new insight into the exhibit. Because it provides the other visitors with a new topic of conversation, we call this category "providing a topic (**Pr**)." Figure 4 (center) shows an example of this category. This contains a lot of text annotations to explain the details of a book.

The third type is "relationship between exhibits (**An**)," which includes link(s) between two or more of the exhibits and explains the relationship between these exhibits. The robot shows these contents in combination with his bodily actions. Figure 4 (right) shows a typical example of the category (**An**). This includes photos taken at several areas and explains the relationships among them using text annotations and arrows as well as the title name of the relationship.

To determine the categories, we assume, by reference to Cone and Kendall [15], that visitors look around a museum with roughly three phases as a standard scenario. In the scenario, the visitors firstly look at an overview of the exhibition (overviewing phase); then, they look closely at a showpiece which they interested to (focusing phase); after looking at several showpieces, they recognize relationships among exhibitions/showpieces (correlating phase). The three categories correspond to the three steps; Short description (Sh), Providing topic (Pr) and Relationship between exhibits (An) corresponding overviewing phase, focusing phase and correlating phase, respectively.

We manually classified the annotated photographs on PhotoChat in the series of experiments and created the content for distribution; however our feasibility study showed that the content classification process can be automated using a support vector machine (SVM). In our feasibility study, we used three features for SVM parameters, the number of pen strokes, the number of hyperlinks and the number of people who made an annotation into the photograph. The number of pen strokes indicate the amount of annotations. This feature can be used as a determinant whether the photograph is informative or not. We assume that the number of hyperlinks classify the class of the photograph. If there are some links and the destination of the link includes the photo taken in another area, the photograph possibly describes a relationship between two different areas. The number of people make annotations into the photograph may indicate the interest level of the photograph. If many visitors made annotations in one photograph, the photograph may interest wide range of people.

This shows the possibilities of automatic classification and distribution of annotated photographs. That is, the robot automatically recognizes the categories of photos and shares them with others. This allows us to realize our concept.

4.3 Design of Robot's Behavior

Figure 5 shows an overview of the robot's behavior. The robot finds a visitor who remains at an exhibit for several minutes, and provides him/her with content accompanied by bodily actions. We set the rules for providing content such that they depend on

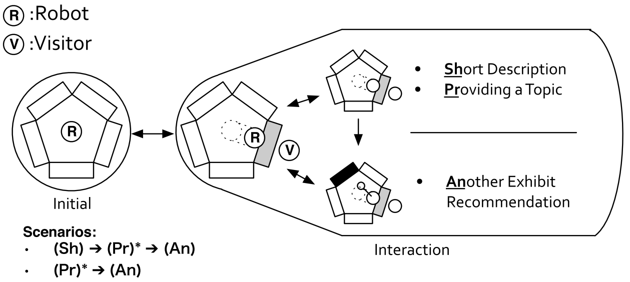


Fig. 5 Overview of the robot's behavior. The robot has three possible courses of action. The first is to provide a short description of the exhibit, the second is to provide a topic about the exhibit, and the third is to recommend another exhibit to the visitor.

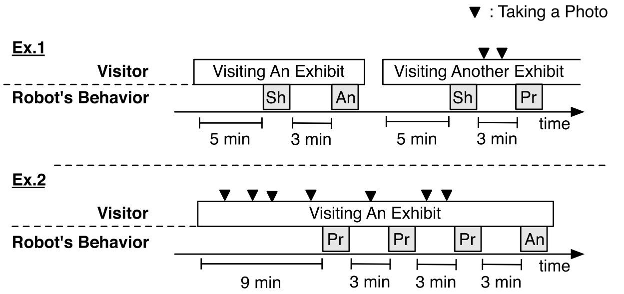


Fig. 6 Two examples of the robot's behavior. The robot varies its behavior depending on the visitor's actions.

Fig. 7 The overview of the two workshops (exhibitions). There are five exhibits in each exhibition (upper). Participants hold chats via PhotoChat (bottom).

the visitor's behavior.

We determined parameters of the rules by conducting a pre-examination. We asked five people to participate in the examination and also asked them to take photographs and to add annotations to the photographs using PhotoChat. The pre-examination took about 50 minutes and was done on the same condition as described in Section 5.1. The delivered rules are described as follows:

- **Sh** (short description)
  The robot provides a "short description" if a visitor meets two conditions: they stay for more than 5 minutes at an exhibit and do not take photographs frequently (less than two photographs per 5 min). The robot states that "This is a short description of this exhibit" while simultaneously distributing "short description" content to PhotoChat. Providing topics will only be related to the exhibit in front of the visitor.
- **Pr** (providing a topic)
  The robot provides a new topic for a visitor if a visitor meets the following two conditions: they stay for over 9 minutes at an exhibit and take photographs frequently (more than three photographs per 9 min). The robot says "You will like this photo because you seem to like the exhibit here" while distributing the appropriate content to PhotoChat. The robot will then continuously provide new topics every 3 minutes if the visitor meets one condition: they have taken at least one photograph in the last 3 minutes. Same as **Sh** condition, providing topics will be about the exhibit the visitor stays.
- **An** (relationship between exhibits)
  The robot recommends another exhibit if a visitor meets two conditions: they have received a "short description" from the robot and have not taken any photographs in the last 3 minutes. The robot states that "There is another exhibit over there," while providing a pointing action and distributing appropriate content to PhotoChat.
  In addition, the robot also recommends another exhibit if there is no distributable content remaining, even if a visitor has continuously shown his/her interest in the exhibit. The robot states "You will like the exhibit over there because you seem to like the exhibit here," accompanied by a pointing action, while distributing appropriate content to PhotoChat.

Figure 6 shows two examples of the robot's behavior. In the upper part of the figure, a visitor first receives a "short description" from the robot because he/she stayed for more than 5 minutes at an exhibit. Then, the visitor received a recommendation for another exhibit because he/she did not take any further action. In this case, the robot guessed that the visitor has no interest in this exhibit. In the lower part of the figure, a visitor continuously received new topics from the robot because the visitor took some pictures during his/her visit. The robot will continuously provide new topics every 3 minutes if a visitor shows his/her interest in the exhibit. In this example, the robot finally recommends another exhibit, because the visitor had not taken a picture in the last 3 minutes.

5. Data Collection: Two Workshops in Exhibition Settings

5.1 Method

We performed two workshops for data collection. As shown in **Fig. 7**, we used a narrow table for each exhibit, with five tables in total. Each exhibit contains about ten books, which are prepared by each exhibitor in accordance with his/her interests (i.e., the theme of the exhibit) beforehand.

Five exhibitors participated in the first workshop. The workshop took approximately 50 minutes. Each participant explained his/her exhibit to the others and they visited each other's exhibits. This means that each participant talked as both an exhibitor and a visitor with the others. We asked all participants to take photographs and to add annotations to the photographs using PhotoChat. Each participant recorded his/her interests and insights into

Table 1 Number of categorized photographs of each area. For the category An, the numbers are overlapped due to it relates several areas.

Area	Sh	Pr	An
1	6	14	3 (1-2, 1-3-4)
2	3	32	2 (1-2)
3	1	16	1 (1-3-4)
4	2	11	1 (1-3-4)
5	3	16	-
sum	15	89	3

the individual books, their findings for each exhibit, and the relationships between the exhibits.

The five exhibitors who had participated in the first workshop and five new visitors participated in the second workshop. The workshop again took approximately 50 minutes. The exhibition setting was same as that of the first workshop. We asked the five visitors to use PhotoChat to take photographs and to add annotations to these photographs. We also asked the exhibitors to explain their respective exhibits. Thus, in this case, the exhibitors did not move away from their exhibits and did not have access to PhotoChat.

5.2 Collected Data

We collected 227 photographs from the series of workshops. We manually classified them into three categories, as described above. In addition to the three categories, we dropped photographs that had no annotation or seemed difficult to categorize. We categorized 15 photos into "short description," 89 photos into "providing a topic," and 3 photos into "relationship between exhibits". This means that 107 out of the 227 photographs are suitable for content distribution. **Table 1** shows the detail number of categorized photographs. We believe that we collected a reasonable number of photos from the series of workshops. However, it might be difficult to find a relationship between areas within the time limitation of the workshop; we had only three photos for the **An** category through the workshop.

We also tried to classify the content using a SVM. In this experimental study, we employed following three features for SVM parameters, the number of pen strokes, the number of hyperlinks and the number of people who annotated into the photograph. In order to evaluate an accuracy of SVM classification, we applied five class cross-validation to the 107 PhotoChat data (i.e., annotated photographs). Result showed that the content classification process can be automated using the SVM with an accuracy rate of 86%.

6. Experiment: Interaction between Visitors and the Robot in Exhibition Setting

We performed a series of experiments to investigate the interactions between the robot guide and the visitors. After the two workshops that were intended to collect data, we created the content for distribution. As explained earlier, we manually classified the PhotoChat data (i.e., the annotated photographs) for distribution. We also implemented the robot behavior, including physical actions such as moving forward and pointing at a distant object. However, because of technical limitations, we did not implement automated physical action functions for the robot. In this case, the movements of the robot were remotely controlled by one of the experimenters. Instead of implementing automated physical action, we took the Wizard of Oz approach (WoZ) [16]. For the sake of helping the experimenter, we developed a notifier program that notifies both the timing of and the instructions for the robot's actions to the experimenter. As we described in the Section 4, the robot provide a content according to the rules except for the third session. We implemented these rules into the notifier program. It notifies the experimenter when the current situation matches the rule.

Eleven participants participated in the series of experiments (three for the first session, five for the second session, and three for the third session). Each of first and second session took about 50 minutes, and third session took about 30 minutes, for a total of 130 minutes. We used two video cameras to study the interactions among the robot and the participants.

To evaluate validity of the rules to provide topics, we changed the rules based on observations through the prior two sessions. More specifically, the robot provides a topic, leaving aside the frequency of taking picture, if a visitor stays for more than five minutes at an exhibit for **Sh** condition and nine minutes for **Pr** condition; we changed each of them to three minutes. This parameter (i.e., three minutes) was determined by the average continuous staying time of each exhibition.

To investigate visitors' acceptance and impressions of the guidance, we also interviewed participants. The interview consisted of roughly two sessions, one was that prearranged and the other was held with video. As for prearranged interview, we asked three questions to each participant; first question asked to evaluate PhotoChat, second one asked to evaluate behaviors of robot and last one asked how did the participant feel about the entire session. As for with video interview, we expect that showing the video will be a clue to remember his/her behaviors.

7. Results: Shared Contents with Robot's Action Affects Human Behavior in Exhibition

During the series of experiments, the visitors freely walked around the space of the exhibition and sometimes browsed a book, took a photograph and added some annotations to the photograph. The robot provided content to the visitors a total of 11 times. **Sh** and **Pr** in the first experiment, **Sh** and **An** in the second one, and four **Sh** and three **Pr** in the last. In the first experiment, the robot provided **Pr** to participant 2 (P2) but P2 ignored it, and also the robot provided **Sh** to P1 and P1 reacted the providing content. In the second experiment, the robot provided **Sh** to P8 but P8 ignored it, however, P6 and P7 were interested in the provided content and moved in front of the robot. P8 also received content **An** and reacted it as moving to exhibition the robot introduced. We will now describe some of our findings from our investigation of the interactions among the robot and the participants.

7.1 Robot Embodiment

When the robot provided content, most of the participants that were spoken to by the robot looked at the robot (nine out of 11 people), and seven of them looked at their PhotoChat screen and

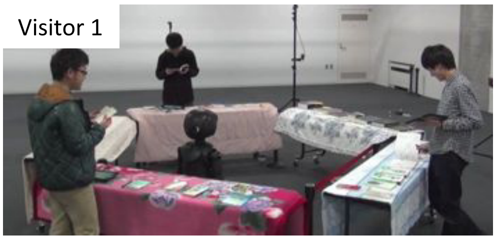
Visitor: browsing books on the exhibit
Robot: approaching to the visitor

Visitor: giving attention to the robot
Robot: “You will like the related information”

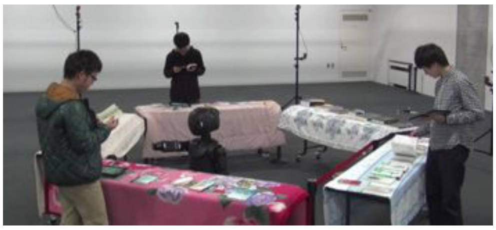
Visitor: checking the provided information on PhotoChat

Visitor: picking up the book shown in the provided information

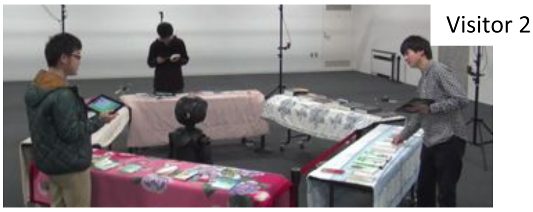
Fig. 8 Typical scene where the robot provided content to a visitor.



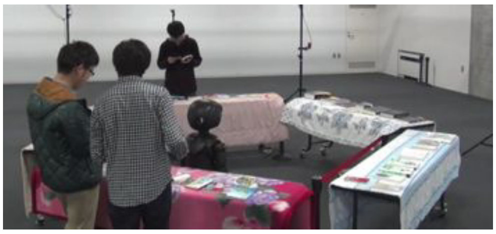
Robot: approaching to the visitor
Visitor1: giving attention to the robot



Robot: “You will like the related information”



Visitor 1: checking the provided information on PhotoChat
Visitor 1: making eye contact with visitor 2



Visitor 2: moving to talk with visitor 1

Fig. 9 Case where the robot’s behavior triggered a real conversation in the space.

confirmed the distributed content. In addition, five people took a book that was related to the distributed content and browsed the book for a little while, as shown in **Fig. 8**.

The robot behavior patterns were only designed to distribute content and to talk to a *single* visitor. However, interestingly, more than half (16 out of 26 [*1]) of the people who were *not* spoken to by the robot also looked at the robot, and 12 of them confirmed the distributed content on PhotoChat. In addition, six out of them (i.e., 12 people) moved to the exhibit for which the robot had provided content, and half of them selected a book to browse. This implies that our robot guidance system successfully affects the visitors’ experiences of an exhibition by distribution of the content collected from visitors who had been there in the past.

We also found an interesting situation where the robot’s behavior triggered a *real conversation* (i.e., not a conversations via PhotoChat) among the visitors in the exhibition space. We observed that the robot’s behavior sometimes drew a visitor who was at another exhibit. In this situation, most of the visitors (five out of six) started a conversation with the others about the exhibit. In contrast to the provision of *personalized* information, because a robot has a real body their presentations can be leaked. This kind of leaked information may then lead to interactions among the people that are present around the space. **Figure 9** shows a typical example of this situation. In this situation, the robot provided information for visitor 1, however, visitor 2 was also interested in the provided information and he moved to talk with visitor 1 about the provided information.

We believe that these findings will not appear without the robot embodiment; In case of conventional (i.e., PhotoChat-only) content distribution, the visitors will not aware where does the content associate with or who does it be addressed to. Some participants addressed the embodiment of the robot. We asked participants, “*Did you understand who is the robot addressed to?*,” with showing a video in the situation that the robot is talking to participant 4 (P4). There’s five people including P4 in the situation and two people, P4 and P8, were in front of the robot. Their answers are follows:

P4: “I noticed that the robot spoke to me.”
P5: “(no answer)”
P6: “I thought that the robot spoke to both [P4] and [P8].”
P7: “[P4].”
P8: “I thought it was for [P4].”

[*1] First session: $2 \times (3 - 1)$ + Second session: $2 \times (5 - 1)$ + Third session: $7 \times (3 - 1)$

This implies that the robot embodiment helped participants to understand who is the robot addressed to, and it could not be happen on conventional PhotoChat system.

Participants also reported, "I don't intend to see photos on PhotoChat because of its inconvenience. (P1)," "It was difficult to follow the timeline of PhotoChat. It was too fast. (P2)" and "I was interested in some of photos but I couldn't find them in the (real) exhibition. (P3)" The robot embodiment may be possible to conquer these problems. P1 may only see photos on PhotoChat when the robot talked to him. P2 may not need to follow the timeline. P3 could see where the robot pointed out and might find where it is.

7.2 Validity of Providing Rules

Through the first two experiments, the robot provided only four clues. This caused because the staying time of each exhibition differed between that of data collection and that of the experiments. We thus investigated the validity of the rules to provide topics by changing the rules based on observations through the first two sessions. Specifically, in the first two experiment, each participant stays an exhibit for three minute in average. We thus changed condition relevant to the staying time. The robot provides a topic, leaving aside the frequency of taking picture, if a visitor stays for more than five minutes at an exhibit for **Sh** condition and nine minutes for **Pr** condition; we changed each of them to three minutes.

Thanks to applying this change, the robot provided seven contents in the third experiment. While the first two experiment took 50 minutes, the last experiment took 30 minutes, 40% shorter. Including the condition, the provided content increased about 3 times (× 2.91).

The suitable providing rules may vary in participants, maybe in contents of the exhibition or in amount of exhibition. This implies it should be adjusted automatically by using average staying time of each exhibition and/or meta data of the exhibition.

7.3 Difference of Distributed Contents

In the experiment, the system distributed 26 contents (photographs) to the visitors. The question is what the difference between the type of distributed contents. We classified contents into three categories (i.e., **Sh**, **Pr** and **An**) and the system provide an appropriate content to a visitor by watching behaviour of the visitor. However, it is unknown how contents affect the behavior of the visitors. Since there are only small sample of data we could collect throughout the experiments, it is difficult to aver the correctness of the sample. However, it may be depends on the state of the visitor. Some of them immediately reacted the distributed content and some of them did not. We ask the reason why did they react to the distributed contents or did not:

P5 (ignored): "I just concentrated to see an exhibit so I could not grasp what the robot said."

P7 (ignored): "I didn't ignored the content. I got a glimpse of it and back to see the current exhibit."

P5 (reacted): "I didn't concentrate anymore ... and I didn't have conversation with my friend. So ... (I reacted to the robot)"

P6 (reacted): "I just finished to see an exhibit and the timing of proving was exactry good to me."

P8 (reacted): "I was interested in the content shown on the PhotoChat and just move to the exhibit space the robot pointed at."

Through the interview, we found that it may depend on the current situation of a visitor whether the visitor react to the content or not, but may not depend on the type of the contents. That is, it is very important to grasp the current situation. Sensors and machine learning technologies may help understanding the situation, for example, we can estimate concentration level by sensing movement of the eyes.

We also found the real conversation may happen when accompanying visitor say to him/herself. P8 said that "I just move to accompanying person because he said that I could not understand the exhibit ..." This kind of soliloquy are frequentry happen after he/she was provided a content from the system and these may play a role as a triger to have a real conversation.

8. Conclusion

In this paper, we have proposed a novel way to share visitor experiences and interests in exhibition spaces and have implemented a system using a robot and PhotoChat. Specifically, we developed a system to collect conversations on PhotoChat and distribute them, accompanied by the robot's bodily actions, to other visitors. We found that our robot guidance system successfully affects the visitors' experiences at an exhibition by distributing content that was collected from the visitors who were there in the past. We also found an interesting situation where the robot's behavior triggered a real conversation among the visitors in the exhibition space.

In this study, we performed workshops and experiments in two phases, where the first was for data collection and the second was to investigate the interactions among the robot and the visitors in the exhibition space. We intend to merge these two phases into a system that will automatically classify the collected data and circulate this data from the past to the future.

We assumed that visitors look around a museum with roughly three phases as standard scenario, and the robot provide topics relied on this standard scenario. We believe that the scenario works on most of cases in museums. However, several cases may not suite the scenario, for example, an interactive exhibition. To adapt interactive showpieces, we may consider a timeline of each of them.

Through the series of experiments, we found some interesting results regarding the robot embodiment, however, there several questions remained. For example, it will be worth to tackle to find the causes and effects of robot embodiment. Also, social acceptance of our system will be another meaningful aspect. Some of visitors may feel annoy because if they do not intend to use the system, the robot talk to someone and all visitor around the robot will be enforced to hear/see the behavior of the robot.

References

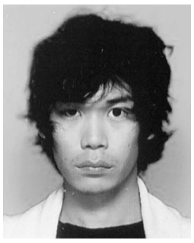
[1] Sumi, Y., Ito, J. and Nishida, T.: PhotoChat: Communication support system based on sharing photos and notes, *CHI 2008 Extended Abstracts*, pp.3237–3242, ACM (2008).

[2] Burgard, W., Cremers, A.B., Fox, D., Hähnel, D., Lakemeyery, G., Schulz, D., Steiner, W. and Thrun, S.: The interactive museum tour-

guide robot, *Proc. 15th National Conference on Artificial Intelligence* (*AAAI-98*), pp.11–18 (1998).
[3] Asoh, H., Vlassis, N., Motomura, Y., Asano, F., Hara, I., Hayamizu, S., Ito, K., Kurita, T., Matsui, T., Bunschoten, R. and Kröse, B.: Jijo-2: An office robot that communicates and learns, *IEEE Intelligent Systems*, Vol.16, No.5, pp.46–55 (2001).
[4] Tezuka, H., Katafuchi, N., Nakamura, Y., Machino, T., Nanjo, Y., Iwaki, S. and Shimokura, K.: Robot platform architecture for information sharing and collaboration among multiple networked robots, *Journal of Robotics and Mechatronics*, Vol.18, No.3, pp.325–332 (2006).
[5] Nakamura, Y., Machino, T., Motegi, M., Iwata, Y., Miyamoto, T., Iwaki, S., Muto, S. and Shimokura, K.: Framework and service allocation for network robot platform and execution of interdependent services, *Robotics and Autonomous Systems*, Vol.56, No.10, pp.831–843 (online), DOI: http://dx.doi.org/10.1016/j.robot.2008.06.008 (2008).
[6] Kanda, T., Hirano, T., Eaton, D. and Ishiguro, H.: Interactive robots as social partners and peer tutors for children: A field trial, *Human-Computer Interaction*, Vol.19, No.1, pp.61–84 (2004).
[7] Shiomi, M., Kanda, T., Ishiguro, H. and Hagita, N.: Interactive humanoid robots for a science museum, *ACM/IEEE 1st Annual Conference on Human-Robot Interaction* (*HRI2006*), pp.305–312 (2006).
[8] Kuno, Y., Sadazuka, K., Kawashima, M., Yamazaki, K., Yamazaki, A. and Kuzuoka, H.: Museum guide robot based on sociological interaction analysis, *Proc. CHI 2007*, pp.1191–1194, ACM (2007).
[9] Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M. and Kuzuoka, H.: Precision timing in human-robot interaction: Coordination of head movement and utterance, *Proc. CHI 2008*, pp.131–140, ACM (2008).
[10] Kuzuoka, H., Pitsch, K., Suzuki, Y., Kawaguchi, I., Yamazaki, K., Yamazaki, A., Kuno, Y., Luff, P. and Heath, C.: Effect of restarts and pauses on achieving a state of mutual orientation between a human and a robot, *Proc. CSCW '08*, pp.201–204, ACM (2008).
[11] Matsumoto, R., Nakayama, H., Harada, T. and Kuniyoshi, Y.: Journalist Robot System: Robot System Making News Articles from Real World, *IEEE/RSJ Internatinal Conference on Intelligent Robotics and Systems* (*IROS 2007*), pp.1234–1241 (2007).
[12] Riether, N., Hegel, F., Wrede, B. and Horstmann, G.: Social Facilitation with Social Robots?, *Proc. 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*, *HRI '12*, pp.41–48, ACM (online), DOI: 10.1145/2157689.2157697 (2012).
[13] Sumi, Y., Mase, K., Müller, C., Iwasawa, S., Ito, S., Takahashi, M., Kumagai, K., Otaka, Y., Tsuchikawa, M., Katagiri, Y. and Nishida, T.: Collage of video and sound for raising the awareness of situated conversations, *Intelligent Media Technology for Communicative Intelligence* (*Second International Workshop, IMTCI 2004, Revised Selected Papers*), LNAI 3490, Bolc, L., Michalewicz, Z. and Nishida, T. (Eds.), pp.185–194, Springer (2005).
[14] Koide, Y., Kanda, T., Sumi, Y., Kogure, K. and Ishiguro, H.: An approach to integrating an interactive guide robot with ubiquitous sensors, *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems* (*IROS 2004*), pp.2500–2505 (2004).
[15] Cone, C.A. and Kendall, K.: Space, Time, and Family Interaction: Visitor Behavior at the Science Museum of Minnesota, *Curator: The Museum Journal*, Vol.21, No.3, pp.245–258 (online), DOI: 10.1111/j.2151-6952.1978.tb00545.x (1978).
[16] Kelley, J.F.: An Iterative Design Methodology for User-friendly Natural Language Office Information Applications, *ACM Trans. Inf. Syst.*, Vol.2, No.1, pp.26–41 (online), DOI: 10.1145/357417.357420 (1984).

Kohei Matsumura is an Assistant Professor at Ritsumeikan University. He received his Ph.D. in Knowledge Science from Japan Advanced Institute of Science and Technology in 2010. His broad research interests include human-computer interaction, embodied cognitive science, urban informatics and novel HCI techniques. He is a member of IPSJ and ACM.

Yasuyuki Sumi is a Professor at Future University Hakodate since 2011. Before joining the university, he has been a researcher at ATR and an Associate Professor at Kyoto University. He received his B.Eng. from Waseda University in 1990, and his M.Eng. and D.Eng. degrees in Information Engineering from the University of Tokyo in 1992 and 1995. His research interests include knowledge-based systems, creativity supporting systems, interface/social agents, ubiquitous/wearable computing, Web intelligence, multimedia processing, and their applications for facilitating human interaction and collaboration.



Takumi Gompei received his M.S. from Future Univercity Hakodate in 2014. During his master program, he was working on several research projects including human-robot interaction and entertainment computing. He is currently working as a system integrator.