Facilitating Experiential Knowledge Sharing through Situated Conversations

Ryo Fujikura
Future University Hakodate
Hokkaido, Japan
r-fujikura@sumilab.org

Yasuyuki Sumi
Future University Hakodate
Hokkaido, Japan
sumi@acm.org

ABSTRACT

This paper proposes a system that facilitates knowledge sharing among people in similar situations by providing audio of past conversations. Our system records all voices of conversations among the users in the specific fields such as tourist spots, museums, digital fabrication studio, etc. and then timely provides users in a similar situation with fragments of the accumulated conversations. For segmenting and retrieving past conversation from vast amounts of captured data, we focus on non-verbal contextual information, i.e., location, attention targets, and hand operations of the conversation participants. All voices of conversation are recorded, without any selection or classification. The delivery of the voices to a user is determined not based on the content of the conversation but on the similarity of situations between the conversation participants and the user. To demonstrate the concept of the proposed system, we performed a series of experiments to observe changes in user behavior due to past conversations related to the situation at the digital fabrication workshop. Since we have not achieved a satisfactory implementation to sense user's situation, we used Wizard of Oz (WOZ) method. That is, the experimenter visually judges the change in the situation of the user and inputs it to the system, and the system automatically provides the users with voices of past conversation corresponding to the situation. Experimental results show that most of the conversations presented when the situation perfectly matches is related to the user's situation, and some of them prompts the user to change their behavior effectively. Interestingly, we could observe that conversations that were done in the same area but not related to the current task also had the effect of expanding the user's knowledge. We also observed a case that although a conversation highly related to the user's situation was timely presented but the user could not utilize the knowledge to solve the problem of the current task. It shows the limitation of our system, i.e., even if a knowledgeable conversation is timely provided, it is useless unless it fits with the user's knowledge level.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*.

KEYWORDS

conversation, situation, experience sharing

ACM Reference Format:
Ryo Fujikura and Yasuyuki Sumi. 2020. Facilitating Experiential Knowledge Sharing through Situated Conversations. In *AHs '20: Augmented Humans International Conference (AHs '20), March 16–17, 2020, Kaiserslautern, Germany.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3384657.3384798

1 INTRODUCTION

This paper proposes a system that facilitates knowledge sharing among people in similar situations by providing audio of past conversations. Our system records all voices of conversations among the users in the specific fields such as tourist spots, museums, digital fabrication studio, etc. and then timely provides users in a similar situation with the related parts of accumulated conversations. Our system focusses on presenting the audio part of the conversations using an earphone or headphones rather than displaying the captured video data on a head-mounted display. This approach gives the possibility to provide the information without disrupting the users, so they can focus on the actual events and overhear potentially interesting topics.

For segmenting and retrieving related conversations from vast amounts of accumulated data, we focus on non-verbal contextual information, i.e., location, attention targets, and hand operations of the conversation participants. All voices of conversation are recorded, without any selection or classification. The delivery of the voices to a user is determined not based on the verbal semantic aspects of the conversation but on the similarity of surrounding situations between the conversation participants and the user. Our basic thought is that the recorded voices may be useful to someone in the future, regardless of the speaker's intent, and that possibility cannot be predicted by anyone in advance. We assume that the matchmaking using contextual information plays sufficient role for facilitating the re-experience of knowledgeable conversations.

AHs '20, March 16–17, 2020, Kaiserslautern, Germany

ACM ISBN 978-1-4503-7603-7/20/03...$15.00
https://doi.org/10.1145/3384657.3384798

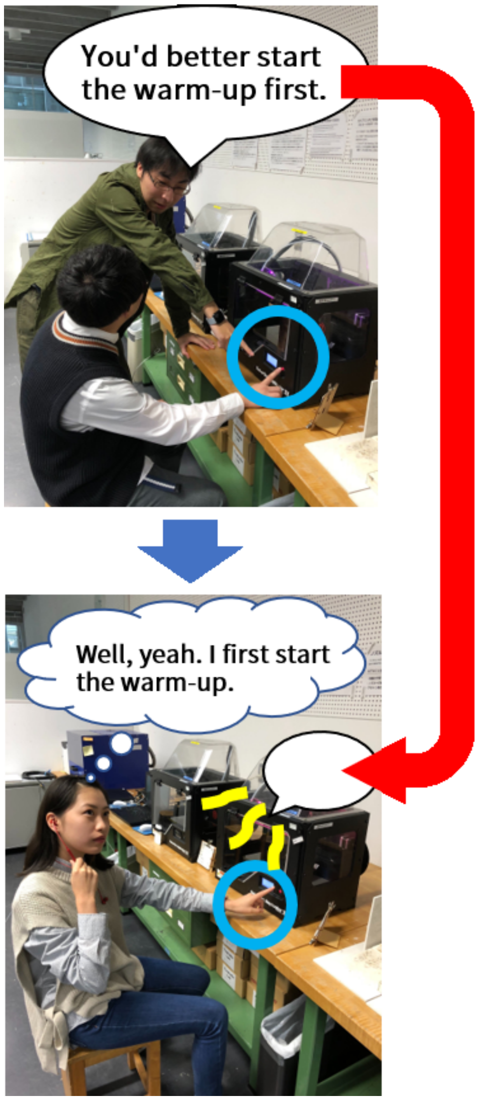

Figure 1: Usage scenario of our system

Figure 1 illustrates an expected scenario of our system used in a digital fabrication studio. The first picture shows a conversation scene during a group work with 3D printer. On the second picture, the voices of the conversation are automatically provided once a novice user is about to use the 3D printer. Retrieving the conversation from a huge amount of recorded conversations is done based on the similarity of contextual information, i.e., the same location (near the 3D printer) and similar behavior (operating the 3D printer).

To demonstrate our concept of the system, we performed a series of experiments to observe changes in user behavior due to past conversations related to the situation at the digital fabrication workshop. Since we have not achieved a satisfactory implementation to sense user's situation, we used Wizard of Oz (WOZ) method. That is, the experimenter visually judges the change in the situation of the user and inputs it to the system, and the system automatically provides the users with voices of past conversation corresponding to the situation.

Experimental results show that most of the conversations presented when the situation perfectly matches is related to the user's situation, and some of them prompts the user to change their behavior effectively. Interestingly, we could observe that conversations that were done in the same area but not related to the current task also had the effect of expanding the user's knowledge. We also observed a case that although a conversation highly related to the user's situation was timely presented but the user could not utilize the knowledge to solve the problem of the current task. It shows the limitation of our system, i.e., even if a knowledgeable conversation is timely provided, it is useless unless it fits with the user's knowledge level.

2 RELATED WORK

There has been many work on wearable computing techniques for recording our daily experiences[10, 23, 24], and extending our memory capabilities[4, 8, 9, 18]. StartleCam can detect "surprise" from the wearer's vital information and capture and record the moment[6]. Forget-me-not is an attempt to understand a user's situation, such as where they met people, and to support memory recall by using the recorded situation for information retrieval[11]. Timemachine computing is a system that records its operations, such as creating, moving, and deleting files, as an index in the desktop environment, and can always return to the former state by referring to it[16].

Nomadic radio[20], wearable spatial conferencing space[2], LiveSphere[15], augment-able reality[17], and Aizawa's life log capture system[1] are examples of research providing information based on real world situations and contexts. These studies explored ways of sharing experiences to promote the discovery and acquisition of knowledge, which is similar to our approach focusing on speech voices.

The following are studies on the recording and analysis of multimodal conversations: NIST[5], AMI[3], VACE[13], and CHIL[25]. NIST and AMI recorded and analyzed conversations in a meeting format using multiple sensors and cameras and constructed the body of the conversation. In contrast, VACE evaluated and analyzed audio information and nonverbal information equally. CHIL attempted to recognize human behavior in conference meeting rooms to support communication by interacting with a computer. In this research, we provide information by presenting conversations; there are already many studies to detect conversations and speakers in conversations[7, 22].

Matsumura et al. proposed a method for recording and presenting conversations in a car to distribute the conversational knowledge among residents in a certain city[12]. Their system collects and distributes the voices of conversation in vehicle by linking with contextual information such as place, time, season, passenger, etc. Our aim in this paper is to generalize their idea from specific domain (car driving) to daily activities, and focuses on bodily scale contexts such as location, attention, and hand operation for dealing with vast amounts of voice data. Saito et al. proposed a conversation quantizer, which is a device for extracting characteristic scenes in a conversation as video[19]. They experimented with extracting conversation fragments from a conference using a conversation quantizer and discussed the features of the extracted conversation fragments. Both Saito et al. and Matsumura et al. are discussing conversation extraction, such as automating retrieval based on the number of times an external input is used. In the conversation group used in the conversation presentation experiment described in this paper, we extracted the parts that appeared to be similar in terms of content based on the author's subjectivity.

Müller et al. proposed a system called Ambient Sound Shower[14] that provides its users with voices of conversations among exhibitors and visitors. The system had three playback modes, i.e., ambient sound mode, exhibit overview mode, and specific conversation mode, which are automatically changes due to the users' non-verbal behaviors such as staying, gazing, etc. DyPERS is a system that presents audio-visual sequences associated with objects for enhancing personal memories[21]. In our research, we classify the spatial situation of speakers and listeners more precisely, focusing on hand operations.

3 PREPARATION OF CONVERSATION CORPUS

In this research, we propose a system to encourage the circulation of knowledgeable conversations by embedding them into situations. Here, we describe the preparation of conversation corpus containing of voices of conversations and their contextual information.

3.1 Recording conversations and their contextual information

We collected conversations for the experiment from users of a digital fabrication workshop, because there are diverse tasks and situations, and conversations are naturally occurring. The workshop uses many devices that require specialized knowledge. For this reason, conversations that encourage discoveries and share knowledge are occurring much more in workshop conversations than in other conversations. Because there are spaces that allow for various tasks to be performed by hand, we can easily connect spoken conversations to those situations.

First, we conducted a workshop to collect conversations and analyze the nature of conversations and situations in the workshop. At the workshop, we recorded data by asking our participants to perform tasks at digital fabrication studio. To elicit conversations related to the task, we recruited participants who have experience handling the equipment and novices. By doing so, there was a natural tendency for the novices to ask and start conversations about the task. The tasks involved making wooden coasters and wooden boxes. Devices and tools used for the tasks include laser cutters, 3D printers, files, drivers, and so on. For recording, we captured first-person video and audio using a wearable camera, GoProHERO4. We conducted the workshop a total of seven times and changed the tasks and the number of people each time. Initially, we conducted the experiment with two people, but we then increased the number of people to three or four and the number of devices used to complicate the task. By increasing the number of people, we were able to increase the number of conversations captured in one preliminary study. When we complicated the content of the tasks, the number of devices used increased, and we were able to increase remarkable situations. As a result, we used the data recorded in the workshop for three sessions where conversations about work were frequent throughout the workshop. Table 1 summarizes the workshops used for the analysis.

3.2 Preliminary analysis of collected data

We analyzed the data recorded in the previous section. First, we browsed the first-person videos and extracted individual conversation scenes. It was necessary to define what a group of conversations was to extract it. Therefore, we subjectively extracted the scenes that we felt were a unit of the conversation content. The length of the extracted conversations varied from 2 seconds to 155 seconds. Looking back at the extracted scenes, we created a corresponding table that recorded the time, location, attention target, and hand operation in detail based on the conversation and the speaker's situation.

We focused on situations that we frequently confirmed through the operation of a prototype to present past conversations developed earlier in workshops. We have been dealing with location as an important cue because we decided that it had a meaningful role in understanding the content of the conversation. When browsing the extracted scenes, it was often found that one conversation is usually between multiple people. We assert that the location of the speaker and the location of all participants is critical. Therefore, we marked the location of all conversation participants for one conversation in the correspondence table. We summarize what we noticed by looking back at the table, extracted scenes, and situations.

- The content of the conversation was often closely related to where it was spoken
- Gaze at devices and tools suggests that participants are talking about them
- Pointing and gestures increased when speaking about an equipment

Based on these facts, we experimented with embedding the conversations recorded in multiple situations and presenting the conversations according to the listener's situation. In the next section, we describe a conversation presentation experiment using the WOZ method.

Table 1: Workshop details

	Duration	Participants	Number of conversations extracted
Workshop 1	1 hour 17 minutes	1 experienced, 1 beginner	94 pieces
Workshop 2	1 hour 6 minutes	2 experienced, 1 beginner	84 pieces
Workshop 3	1 hour 9 minutes	3 experienced, 1 beginner	126 pieces

Table 2: Samples of collected conversation data

yyyy/MM/dd H:mm:ss	Duration(s)	Location	Attention target	Hand operation
...	...	...	...	...
2019/6/20 10:52:38	33	Near laser cutter PC	laser cutter PC	Pointing at PC
2019/6/20 10:53:15	40	Near laser cutter PC	laser cutter PC, MDF	Pointing at PC, gesture
2019/6/20 10:53:54	81	Near laser cutter PC	laser cutter PC	Pointing at PC, mouse operation
...	...	...	...	...

4 EXPERIMENTAL EVALUATION OF OUR SYSTEM: EFFECT OF CONVERSATIONS WITHIN THE SAME TASK

To confirm the usefulness of presenting past conversations according to the situation, we conducted a conversation presentation experiment using data collected at the workshop. In this experiment, we instructed the participants to perform the task while listening to the conversations in the workshop. The purpose of this experiment was to confirm the effect of being presented with the conversation, so we did not use the technology to detect the situation automatically, but we experimented using the WOZ method this time. We chose what conversation to present from the 304 conversations according to the listener's situation. In the following, we show the experimental preparation, outline of the experiment, results, and discussion.

4.1 Preparation

We show the specific experimental procedure. First, we predicted that the subject's situation would change rapidly, so we thought that a system to select the conversation according to the subject's instantaneous situation was necessary. Therefore, we implemented a system that meets the above requirements. We implemented it using JavaScript and developed a simple button-type conversation presentation system that can be executed in a local environment. The role of this system is the following two.

(1) Select and manage conversations stored in the database according to the pressed button
(2) Replay selected conversation

We were always present with participants, checked the status of participants visually, and typed into the system so that participants heard conversations selected by it from the wireless earphone. In our experiments, we deal with three factors: location, attention target, and hand operation. The reason for choosing these is discussed in the previous section.

Table 3: The number of conversation data corresponding to each contextual information

Location	(pieces)
3D printer	94
Workbench	83
Laser cutter PC	67
Laser cutter	56
Central area	12
Laser cutter room	8
Tool shelf	1

Attention target	(pieces)
Laser cutter PC	66
3D printer	49
MDF	34
Laser cutter	26
3D printer PC	26
3D model	26
Tool	11

Hand operation	(pieces)
Mouse operation	68
Pointing	66
Gesture	63
Holding MDF	24
Button operation	21
Bonding	17
Polishing	12
Screwing	11

We show the number of conversations that correspond to each situation item, shown in Table 3.

Another option for the situation to focus on was a mutual gaze on a subject. We excluded this situation from the discussion because a joint gaze accompanied most of the conversations at the workshop. As mentioned in the previous section, we believe that not only the situation of the speaker but also that of everyone who participated in the conversation is critical. Therefore, in this system, we linked conversations stored in the database with situations of all the people involved in conversations. Additionally, we have set the conversation to not play unless the participant's location matches one of the locations in the system. In other words, no conversation is presented unless all three factors match (e.g. location, attention target, hand operation).

Additionally, as shown above, during the experiment, it is considered that there are instances where the subject's situation changes rapidly. However, if different conversations are played each time the situation changes, it will not be possible to hear a single conversation. For this reason, we created an algorithm of the system to play conversations according to situations if the situation lasts for 5 seconds. 5 seconds are seconds that we subjectively feel just right when we used the system in the workshop. For example, a conversation is presented if we apply only location to the system for 5 seconds.

Even when all of three aspects of contextual information were specified for the current situation, there were still several numbers of past conversations with the same situation, so we frequently had the problem of not being able to select to only one conversation to present. In these cases, we decided which conversation would be presented at random.

We did not perform a simultaneous presentation of conversations in this experiment because we wanted to maintain the intelligibility of conversations. One of the reasons is that in the room where the laser cutter is located, the environmental sounds, such as the ventilation fan, are troublesome.

In one experiment, there was initially a choice to not repeat a conversation that had already been played. However, to verify the effect of missing or listening to a conversation multiple times, we made it possible to select the conversations that had already been played. We conducted experiments using a system that implements an algorithm that satisfies the above constraints.

4.2 Experiment outline

We show the experiment in Figure 2. We conducted three experiments with one pair of university students. Any participant who participated in the experiment once did not participate in other experiments. In order to observe the effect of the conversation during work in digital fabrication studio, we asked the pair of participants to facilitate verbally externalizing their thoughts during the task. All pairs were close enough to exchange words regularly. During the experiment, one of the two had wireless earphones to listen to the conversations. We did not instruct the earphone wearers to listen carefully to the presented conversation. We set

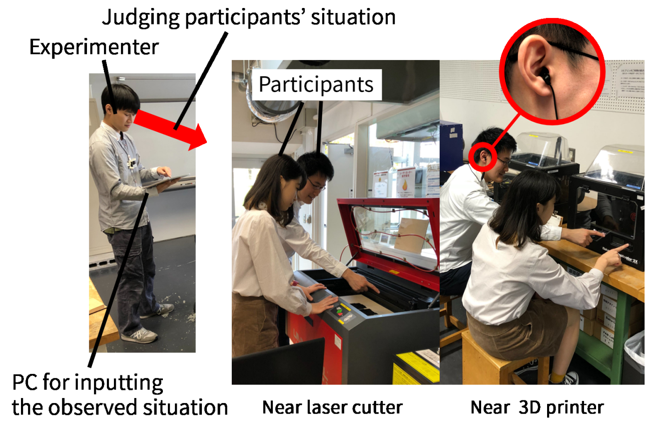

Figure 2: Experiment for system evaluation

up an experimental setting in which they would pay attention if they were interested in the presented conversation. We thought that wearing earphones in both ears would hinder conversation between participants, so we instructed them to wear only one ear bud. We show the attributes of the experiment's participants in the Table 4.

We recorded first-person images during the experiment by having all participants wear a GoProHERO4. In the experiment, we gave the participant the task of making a wooden box, as in the workshop. The wooden box was mainly made of 6 mm thick MDF, ABS resin, and screws. Equipment and tools used in the experiment include laser cutters, 3D printers, files, drivers, and bonds. Each experiment took about one and a half hours.

The experimenter judged the location, attention target, and hand operation around the participant wearing the earphones and inputted manually them into the system. We show an example of how a participant is presented with a conversation in Figure 3. For example, conversations selected by contextual information such as a location near a 3D printer, the direction of the 3D printer, and the button operation are presented to the participants, shown in Figure 3. Since there are many candidates only for location matching like the 3D printer, such as 94 pieces, so an out-of-focus conversation is often presented. However, conversations involving broad knowledge are also likely to be presented. On the other hand, when conversations are matched according to detailed situations such as attention target and hand operation, candidates are selected to specific conversations. If the candidates are not selected to one conversation, one of them is randomly selected. We expected that most conversations would be enough applicable to the specific situation in such cases.

The purpose of the experiment was to confirm the following.

- Can participants understand the content of the presented conversations?

Table 4: Participant attributes

	Participant	Gender	Age	Laser cutter experience	3D printer experience	Earphone attached
Pair 1	participant A	male	23	yes	no	yes
	participant B	male	22	no	yes	no
Pair 2	participant C	male	24	yes	yes	yes
	participant D	male	22	yes	no	no
Pair 3	participant E	male	22	no	yes	yes
	participant F	female	22	yes	no	no

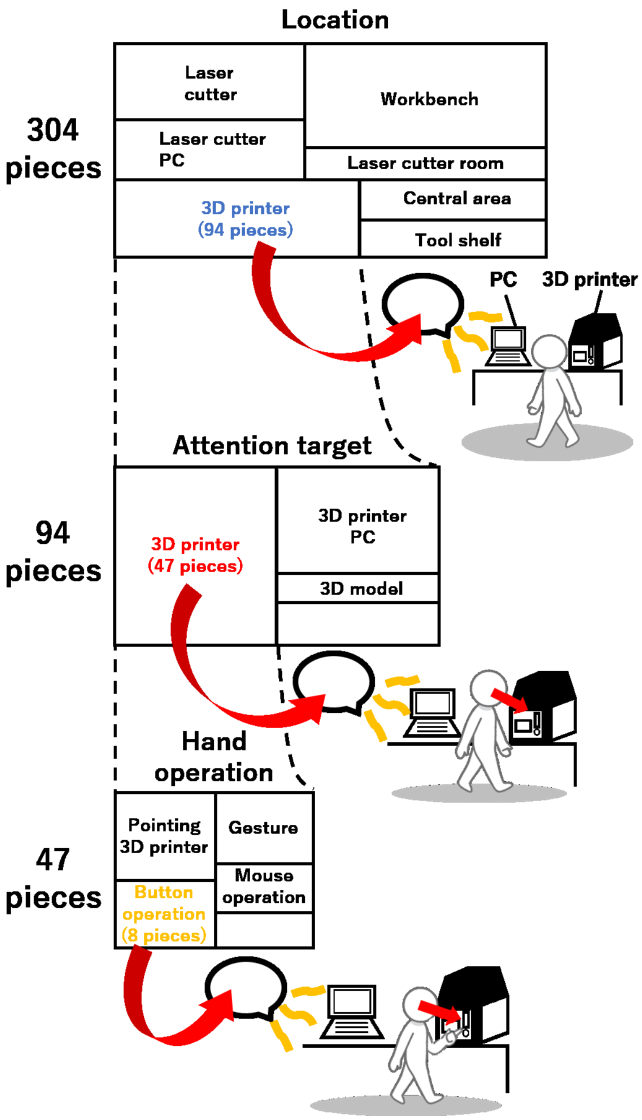

Figure 3: Example scenario of focusing conversation candidates along specific situations

- Are participants influenced by the presented conversations?
- Does new conversations among the participants start in response to the presented conversations?
- Do the conversations provide timely useful information?

After the experiment, we interviewed participants with simple questions. The authors examined the video recorded by the GoProHERO4 and carefully checked the responses and actions of the subjects and the conversations presented by the system. Therefore, subjects were instructed in advance to speak alone and react as much as possible if they were interested in the presented conversation. The next section describes the experimental results and considerations based on them.

4.3 Results and discussion

By viewing the recorded video and checking the log of the presented conversations, we confirmed what kind of conversation was presented and what kind of reaction participants responded with. Here, as shown in Figure 4, we explain the process from hearing the conversation to obtaining the knowledge. First, we mention that the participant listens to the presented conversation. Even if a conversation is presented, if the participants do not hear it, this is an obstacle to acquiring knowledge. Next, even if participants listen to the conversation, if they are not interested in it, they will not gain knowledge. We think the second step is to map the content of the conversation to the situation. Lastly, we list the steps to understand the content of the conversation and gain knowledge. We believe that by clearing these gaps (perceptive, contextual, cognitive), participants can gain knowledge from the conversation. Conversely, if you do not go through the above process, participants will not be able to gain knowledge. From here, we introduce examples where we gained knowledge from conversations and examples where we did not, using conversations in experiments as subjects.

4.3.1 *Effect of replayed conversation.*

We explain the effects of replayed conversations using actual conversations as examples. We focused mainly on times where presenting a conversation changed the participant's behavior. We used the change of behaviors as a clue that participants gained knowledge through the conversations they heard. What follows is a conversation that occurred when trying to change the filament used in a 3D printer during the first experiment. Participant B has forgotten how to replace the filament.

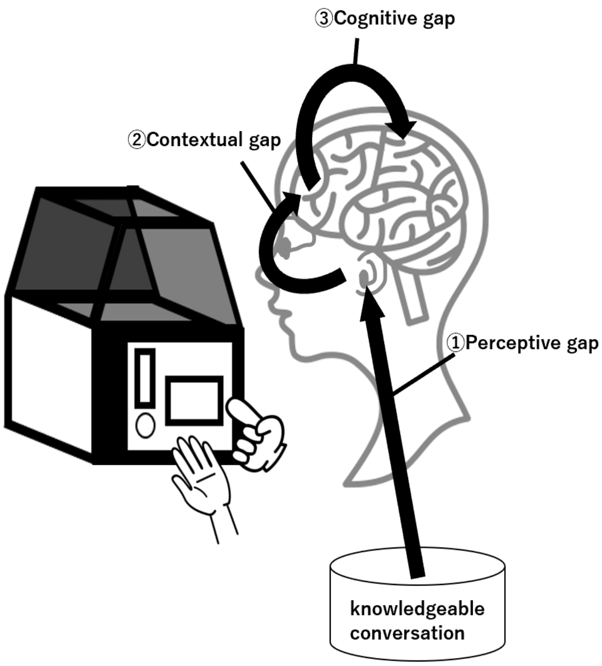

Figure 4: Three difficulties towards effective use of knowledgeable conversations

Participant A *"According to the voice, you have to press the button change filament."*
Participant B *"Change filament?"*
Participant A *"Yes."*
Participant B *"Okay, I found the change filament button."*

The conversation started, and participant B operated the button to select the change filament item. The conversation presented in this way affected the work. The conversations presented in the case was as follows. Speaker A is a person who worked in the past while talking.

Speaker A *"Which should I choose for changing the filament? ... Setting? Preheat? Seems not... This is it? Ah, yes. I found the buttton for change filament."*

There were some other examples in which participants gained knowledge by presenting conversations. The following example shows that the conversation provided by the system induced beneficial conversation between the participants. Participants E and F were chatting near the 3D printer about how to remove the 3D model during the third experiment.

Participant E *"Have you ever removed (the product from the 3D printer table) by yourself?"*
Participant F *"Yes, but I was helped by Ms. N (a manager of the digital fabrication studio) at that time."*
...
Participant E *"The work (of removing the product from the table) should be done sooner."*
Participant F *"I see, it becomes difficult to remove when it cools down."*

At this time, participant E was listening to the following conversation provided by the system.

The studio manager *"The time it's completed is important. You can remove it easier if you do when it's still hot, but it becomes harder after being cooled."*

Here, the presented conversation provided timely and helpful knowledge to participant F for removing the product from the 3D printer table easier. Additionally, he started a conversation with participant F to share it with participant E.

As seen in the example, the voices of conversation presented by the system caused new conversation among the participants, and such conversation has potential to externalize hidden knowledge among them and deepen their understand. In that sense, recording conversations inspired by the voices provided by the system and utilizing them for future knowledge circulation would be useful as clues for refining and highlighting certain knowledgeable conversations. It is a promising future direction of our work.

4.3.2 Location dependence of system effects.

We can say the participants could generally receive the meaningful information from the voices of conversation provided by our system. All three participants said, in the interviews after the experiment, that they could understand the presented conversations. In fact, we often observed the participants responded to the replayed conversation saying "I see."

Curiously, the response to the voices of conversation provided by our system around the 3D printer was much more observed than elsewhere, despite they worked in three places, i.e., laser cutter, 3D printer, and workbench. Regarding the number of conversations in the database, out of 304 total conversations, 123 conversations were presented near the workbench, 94 were presented near the 3D printer, and 83 were presented near the workbench. We cannot see significant difference regarding the number of provided voices of conversation: 123 conversations were presented near the laser cutter, 94 near the 3D printer, and 83 near the workbench.

The main reason why the system was less effective around the laser cutter was that the sound of the laser cutter was too loud to hear the voices provided by the system: That means the participants could not pass the "perceptive gap" mentioned in Figure 4. Another reason may be due to the amount of work done around the equipments. The participants had more work to do around the workbench and laser cutter. Whereas they had more spare time around the 3D printer and just spent time observing the 3D printer. It might provide the participants with the opportunity to carefully listen to the voices presented by the system around the 3D printer and to converse initiated by them.

4.3.3 System limitations.

Here, we describe the limitations of presenting conversations in this system that we felt through experiments. First, the amount of information is small. For example, suppose that the presented conversation alerts the listener saying “I need the nipper now.” However, the listener cannot understand what the nipper is, or where the nipper is, if they do not already know. This is an example of a time the “cognitive gap” was not cleared. If this was a visual information presentation, you could judge which object is the nipper from the appearance of the nipper. This is an unavoidable problem when presenting audio information. However, in the workshop, because there are many necessary situations at hand, the presentation of voice information is more suitable than a visual approach.

As we have seen, our system is sensitive to change in situations of users. Therefore, when five seconds elapse after the situation changes, even if an important conversation is played back, it is interrupted and another conversation is selected to play back. It may cause a trouble. For example, suppose the following voices of conversation is being played back.

> *“Then, let's turn off them. The earlier we turn off, the sooner we can finish.*
> *Be careful that turning off the PC is last. You must keep the order of turning off: the 3D printer first, and the PC last.*
> *You must not turn off the PC earlier (than the 3D printer).*
> *Incorrect order may lead to equipment failure.”*

There is no problem if the users listen to the conversations to the end. However, if they listen only the first part of the conversation, they may turn off the PC even though he should turn off the 3D printer first. This is an example which participants cannot get correct knowledge from the conversation due to “cognitive gap”. The system should be designed so that the entire conversation is played to prevent such troubles.

Finally, we discuss the inhibition of communication between collaborators by using our system. In this experiment, one of the pair participants were asked to listen to the conversation by attaching the wireless earphone to one ear. Therefore, if she/he concentrates on the conversations being played back, she/he may not hear the voices of the other participant, which may hinder face-to-face communication between them. Another problem with providing the voices using earphone was that the other participant could not know if the participant with the earphone was listening something or not. To solve these problems, we decided to use neck speakers instead of earphones in the next experiment. By using the neck speaker, the voices presented from the system can be heard not only by the system user but also by the other participant in the pair, and the system user does not be separated from surrounding sounds.

5 EXPERIMENTAL EVALUATION OF OUR SYSTEM: EFFECT OF CONVERSATIONS FROM DIFFERENT TASK

In the conversation presentation experiment described in the previous section, the task performed by the participant and the task performed when the presented conversation was recorded were the same. We performed another WOZ experiment to verify if presenting conversations was effective even when performing another task. From now on, the conversation presentation experiment described in the previous section is called experiment 1, and the experiment performed again is called experiment 2.

5.1 Experiment

For the experiment 2, We chose a workshop to make a toy (maze), which differs from the process of making a wooden box. The equipment and tools used for the experiment include laser cutters, 3D printers, and bonds. The maze was made of 2.5 mm and 6 mm thick MDF, 2 mm acrylic board, and ABS resin. By preparing multiple plates of different thicknesses and materials, the task of focusing on the laser cutter operation was made more complicated than in experiment 1. On the other hand, the work of sanding and using the screwdriver was not included in experiment 2. Consequently, the conversations related to polishing and the screwdriver were designed to be able to verify whether the listener would consider them noise or turn them into useful conversations. Each experiment took about one hour.

The experiment was performed three times on a pair of university students who participated in experiment 1. The reason for employing the same participants is that we wanted to minimize individual differences in responses to conversation. During the experiment, a person wearing wireless earphones in experiment 1 was asked to wear them in the same way for experiment 2.

We show the attributes of the experiment participants in the Table 4

GoProHERO4 was attached to all the contributors. In the experiment, participants were tasked with creating a maze. We judged the situation around the participants wearing earphones and selected and played back the conversation by manually driving the system. After the experiment, we conducted interviews with simple questions and performed the analysis in the same manner as in Experiment 1. In the next section, we describe the experimental results and considerations based on them.

5.2 Results and discussion

We summarize the results of experiments 1 and 2 in Table 5. In the following, we describe the effect of playing conversations and a comparison of the results of the experiments.

5.2.1 Effect of replayed conversation.

In experiment 2, there was an example in which the behavior

Table 5: Summary of experimental results

		Conversations presented (pieces)	Conversations reacted (pieces)	Conversations that affected behavior (pieces)
Experiment 1 (same task)	Pair 1	220	26	2
	Pair 2	152	31	1
	Pair 3	139	8	2
Experiment 2 (different task)	Pair 1	121	9	0
	Pair 2	178	26	1
	Pair 3	139	9	0

was changed by the conversation presented. During the second experiment, Participant C was watching the operation of the device in front of the 3D printer. Then, he muttered, *"I see ... the first layer!"* In response to this scene, we asked him what he had learned at the time of the hearing. He then said that a timely conversation was played, and he learned from the conversation when to leave the 3D printer. When printing a model with a 3D printer, there is a rule stating that, after starting a print, from the output of the raft to the second layer, we must observe the operation of the device for safety. The conversation gave knowledge of the rule to Participant C, who had forgotten it. Incidentally, the conversation presented here is as follows.

Speaker A *"We have to keep our eyes on the first few layers to see if the 3D printer is working properly.*
Speaker B *"I got it."*

In this way, it seems easier to timely share knowledge through past conversations about devices such as 3D printers that work similar even if the products are different. The above conversation was presented when the location "3D printer area" and attention target "3D printer" were given to the system. As shown in Table 5, we observed a total of six "conversations that affected behavior" through two experiments. One conversation was selected by three situations, four were selected by two situations, and one was selected by one situation. From these results, we can see that just selecting conversations based on location and attention target can provide useful conversations. Although we acknowledge that results are small in number, it seems to me that it is important to observe and analyze many of these cases in daily operations in the future.

However, as described in Table 5, we could not observe another case in which the conversations presented by the system affected the participants' behavior except for the above case. There are several reasons for this, as described below.

- Many of the past conversations presented by the system were not applicable to the situations in the different task.
- Because we recruited the same participants for both of experiments 1 and 2, so they got used to conversations presented by the system.
- The workshop time for the experiment 2 was too short.

5.2.2 Other findings.
We describe here interesting cases, despite they are different from our expectations.

The first case is that Participant E working at the workbench suddenly started polishing the parts for the maze after saying *"I have to polish..."* to himself, despite the polishing was not necessary in the workshop. According to the interview after the workshop, he was induced to do it because the polishing sound around the workbench was provided by the system. We found the episode interesting because it indicates that not only conversations but also surrounding sounds have potential to change a user's behaviors.

In another case, Participant F was in trouble to use the laser cutter because she temporarily forgot how to initialize the laser cutter. Beside her, Participant E was listening the voices of conversations provided by the system but could not help her. We, the experimenters, felt regrettable that the system could not provide effective conversations related to the situation. The fact, however, was not really: We realized that our system timely provided Participant E with good conversations about the laser cutter initialization by checking the system log after the workshop. The reason why Participant E could not deliver the knowledge gained by the provided conversations to Participant F is that Participant E could not understand what was spoken in the provided conversation because of the lack of Participant E's knowledge on laser cutter. It means Participant E could not overcome either of "contextual gap" and "cognitive gap" in Figure 4 because of the lack of knowledge on the facing situation.

6 CONCLUSIONS

In this paper, we proposed a system that facilitates knowledge sharing among people in similar situations by providing audio of past conversations. For segmenting and retrieving past conversation from vast amounts of captured data, we focused on non-verbal contextual information, i.e., location, attention targets, and hand operations of the conversation participants. To demonstrate the concept of the proposed system, we performed a series of experiments to observe changes in user behavior due to past conversations related to the situation at the digital fabrication workshop.

Experimental results showed that most of the conversations presented when the situation perfectly matches is related to the user's situation, and some of them prompts the user to change their behavior effectively. We also observed a case that although a conversation highly related to the user's situation was timely presented but the user could not utilize the knowledge to solve the problem of the current task. It shows the limitation of our system, i.e., even if a knowledgeable conversation is timely provided, it is useless unless it fits with the user's knowledge level.

REFERENCES

[1] Kiyoharu Aizawa, Tetsuro Hori, Shinya Kawasaki, and Takayuki Ishikawa. 2004. Capture and efficient retrieval of life log. In *Pervasive 2004 workshop on memory and sharing experiences*. 15–20.

[2] Mark Billinghurst, Jerry Bowskill, Mark Jessop, and Jason Morphett. 1998. A wearable spatial conferencing space. In *Second IEEE International Symposium on Wearable Computers (ISWC ' 98)*. 76–83.

[3] Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, et al. 2005. The AMI meeting corpus: A pre-announcement. In *International Workshop on Machine Learning for Multimodal Interaction*. 28–39.

[4] Richard W. DeVaul, Alex "Sandy" Pentland, and Vicka R. Corey. 2003. The memory glasses: Subliminal vs. overt memory support with imperfect information. In *7th IEEE International Symposium on Wearable Computers (ISWC ' 03)*. 146.

[5] John S Garofolo, Christophe Laprun, Martial Michel, Vincent M Stanford, and Elham Tabassi. 2004. The NIST meeting room pilot corpus. In *Fouth International Conference on Language Resources and Evaluation*.

[6] Jennifer Healey and Rosalind W Picard. 1998. StartleCam: A Cybernetic Wearable Camera. In *2nd IEEE International Symposium on Wearable Computers (ISWC ' 98)*. 42–49.

[7] Noboru Kanedera, Asuka Sumida, Takao Ikehata, and Tetsuo Funada. 2006. Subtopic segmentation in lecture speech for the creation of lecture video contents. 37, 10 (2006), 13–21.

[8] Takekazu Kato, Takeshi Kurata, and Katsuhiko Sakaue. 2002. Face registration using wearable active vision systems for augmented memory. In *Digital Image Computing: Techniques and Applications*. 252–257.

[9] Tatsuyuki Kawamura, Yasuyuki Kono, and Masatsugu Kidode. 2002. Wearable interfaces for a video diary: Towards memory retrieval, exchange, and transportation. In *6th IEEE International Symposium on Wearable Computers (ISWC ' 02)*. 31.

[10] Nicky Kern, Bernt Schiele, Holger Junker, Paul Lukowicz, and Gerhard Troster. 2002. Wearable sensing to annotate meeting recordings. In *Sixth International Symposium on Wearable Computers (ISWC '02)*. 186–193.

[11] Mik Lamming and Mike Flynn. 1994. Forget-me-not: Intimate computing in support of human memory. In *FRIEND21: International Symposium on Next Generation Human Interface*. 150–158.

[12] Kohei Matsumura and Yasuyuki Sumi. 2014. What are you talking about while driving?: An analysis of in-car conversations aimed at conversation sharing. In *6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI ' 14)*. 8.

[13] David McNeill. 2005. Gesture, gaze, and ground. In *Second International Conference on Machine Learning for Multimodal Interaction*. 1–14.

[14] Christof E. Müller, Yasuyuki Sumi, Kenji Mase, and Megumu Tsuchikawa. 2004. Experience sharing by retrieving captured conversations using non-verbal features. In *1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE' 04)*. 93–98.

[15] Shohei Nagai, Shunichi Kasahara, and Jun Rekimoto. 2015. LiveSphere: Sharing the surrounding visual environment for immersive experience in remote collaboration. In *Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI ' 15)*. 113–116.

[16] Jun Rekimoto. 1999. Time-machine computing: A time-centric approach for the information environment. In *12th Annual ACM Symposium on User Interface Software and Technology (UIST ' 99)*. 45–54.

[17] Jun Rekimoto, Yuji Ayatsuka, and Kazuteru Hayashi. 1998. Augment-able reality: Situated communication through physical and digital spaces. In *2nd IEEE International Symposium on Wearable Computers (ISWC ' 98)*. 68.

[18] Bradley J. Rhodes. 1997. The Wearable Remembrance Agent: A System for Augmented Memory. In *1st IEEE International Symposium on Wearable Computers (ISWC ' 97)*. 123.

[19] Ken Saito, Hidekazu Kubota, Yasuyuki Sumi, and Toyoaki Nishida. 2005. Support for content creation using conversation quanta. In *2005 International Conference on New Frontiers in Artificial Intelligence*. 29–40.

[20] Nitin Sawhney and Chris Schmandt. 2000. Nomadic radio: Speech and audio interaction for contextual messaging in nomadic environments. *ACM Transactions on Computer-Human Interaction* 7, 3 (Sept. 2000), 353–383.

[21] Bernt Schiele, Nuria Oliver, Tony Jebara, and Alex Pentland. 1999. An interactive computer vision system DyPERS: Dynamic personal enhanced reality system. In *First International Conference on Computer Vision Systems (ICVS ' 99)*. 51–65.

[22] Tanja Schultz, Alex Waibel, Michael Bett, Florian Metze, Yue Pan, Klaus Ries, Thomas Schaaf, Hagen Soltau, Martin Westphal, Hua Yu, and Klaus Zechner. 2001. The ISL meeting room system. In *Workshop on Hands-Free Speech Communication (HSC-2001)*.

[23] Takamasa Ueda, Toshiyuki Amagasa, Masatoshi Yoshikawa, and Shunsuke Uemura. 2002. A System for retrieval and digest creation of video data based on geographic objects. In *13th International Conference on Database and Expert Systems Applications (DEXA ' 02)*. 768–778.

[24] Ryoko Ueoka, Koichi Hirota, and Michitaka Hirose. 2001. Wearable computer for experience recording. In *11th international conference on artificial reality and telexistence (ICAT ' 01)*.

[25] Alex Waibel, Hartwig Steusloff, Rainer Stiefelhagen, and Kym Watson. 2009. Computers in the human interaction loop. In *Computers in the Human Interaction Loop*. Springer, 3–6.