Acquiring Surrounding Visual Information Without Actively Taking Photos for People with Visual Impairment

Masakazu Iwamura[1](✉), Takaaki Kawai[2], Keigo Takashima[2], Kazunori Minatani[3], and Koichi Kise[1]

[1] Graduate School of Informatics, Osaka Metropolitan University, Sakai, Japan
{masa.i,kise}@omu.ac.jp

[2] Graduate School of Engineering, Osaka Prefecture University, Sakai, Japan
sbb01097@st.osakafu-u.ac.jp

[3] The National Center for University Entrance Examinations, Tokyo, Japan
minatani@rd.dnc.ac.jp

Abstract. Recent advancements in recognition technology allow people with visual impairment to obtain visual information from their surroundings using smartphone apps and assistive devices. This paper points out a problem with this approach that has not attracted much attention. That is, the user is required to actively take a photo, which is not always easy for people with visual impairment. To address this problem, in contrast to the current standard approach, which we call active information acquisition, we propose passive information acquisition (PIA), which does not require the user to actively take a photo. However, PIA creates a new problem: the app tends to transfer too much information to the user. Therefore, this paper explores better ways for people with visual impairment toward obtaining only the desired visual information in PIA. Specifically, we experimented with nine people with visual impairment to evaluate seven information transmission methods, including information summarization and interactive communication methods.

Keywords: People with visual impairment · Object detection · Summarization · Information selection · Passive information acquisition

1 Introduction

Recognition technology has been employed to function as an eye for people with visual impairment in order to obtain visual information from their surroundings (e.g., [8,12–14]). It has been installed on smartphone apps (e.g., Seeing AI, Envision AI, and TapTapSee) and assistive devices (e.g., OrCam MyEye2 and Envision Glasses) as an indispensable tool. The tacit preconditions of using such apps and devices, however, require the user to actively take a photo using three steps:

K. Miesenberger et al. (Eds.): ICCHP-AAATE 2022, LNCS 13341, pp. 229–240, 2022.
https://doi.org/10.1007/978-3-031-08648-9_27

they must (1) notice a target object or textual information, (2) know its location, (3) take a picture by aiming their camera at that location. We call this standard framework *active information acquisition (AIA)*. A fundamental question that arises here is whether people with visual impairment can notice information around them and determine its location. Even if they succeed in determining the location, it is easy to predict that they will often face the difficulty of taking an appropriate picture for the recognition technology [4,7,9,11,16,19]. Thus, actively taking a photo is not always possible for people with visual impairment. However, this significant problem has not attracted much attention.

This paper focuses on the issue and introduces a new framework that enables the user to acquire the visual information from the user's surroundings without actively taking a photo. In contrast to the conventional AIA framework, we call the new framework *passive information acquisition (PIA)*. A possible method to actualize the PIA framework is to constantly photograph the user's surroundings and recognize all of them; constant recording by a camera (preferably, a wide field-of-view camera) can capture the user's surroundings to which the user is not easily able to pay attention. As a consequence, however, the user obtains a huge amount of surrounding visual information, which is substantially larger than the amount of those that can be processed by a conventional AIA framework. If such a large amount of information were constantly described, the user would be overwhelmed. Let us take the concrete example of object detection. Suppose that ten objects are detected every second through the constant recording and recognition; after one minute, 600 objects have been detected. This amount of information is too much to describe. However, it is likely that the same objects are detected multiple times. Therefore, if such duplication is suppressed by a summarization technique (say, *temporal summarization*), the amount of information conveyed to the user is greatly reduced. Even if the objects are not the same, some could be categorized into the same category, such as "drink bottles." In this case, instead of describing the name of each product, saying "drink bottles" would be more concise. We call this approach *semantic summarization.* These examples indicate that information summarization is crucial in the PIA framework. In this paper, we implemented the naive method (without using information summarization) and three information summarization methods, as shown in Fig. 1.

The discussion above opens the door to another problem that has been overlooked: when using the apps and devices, the user must keep listening to the voice verbally describing the recognition results, even though most of them are not meaningful to the user. Let us imagine that ten objects have been detected. Then, the user needs to evaluate the information about each of the ten objects one by one. In contrast to the behavior of sighted people, who can focus on the information of interest without evaluating everything in their sight, the current approach of the apps is far from efficient. A possible solution to this problem is to introduce *interactive communication*; the user requests information from the app, and then the app tells the requested information to the user. We use a question answering (QA) system to realize it. We prepared three interactive com-

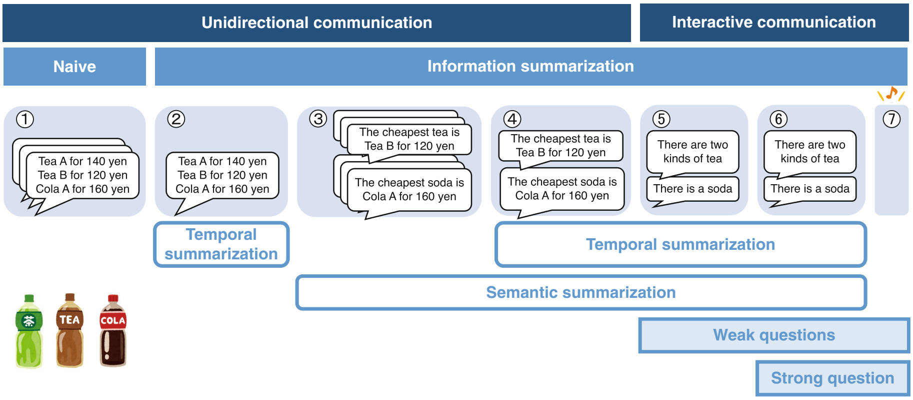
Unidirectional communication
Interactive communication
Naive
Information summarization
①
Tea A for 140 yen
Tea B for 120 yen
Cola A for 160 yen
②
Tea A for 140 yen
Tea B for 120 yen
Cola A for 160 yen
③
The cheapest tea is Tea B for 120 yen
The cheapest soda is Cola A for 160 yen
④
The cheapest tea is Tea B for 120 yen
The cheapest soda is Cola A for 160 yen
⑤
There are two kinds of tea
There is a soda
⑥
There are two kinds of tea
There is a soda
⑦
Temporal summarization
Temporal summarization
Semantic summarization
Weak questions
Strong question
TEA
COLA

Fig. 1. Overview of the seven information transmission methods evaluated in the user study. Four unidirectional communication methods were designed to examine all four combinations of with/without temporal/semantic summarization. Three interactive communication methods are primarily compared with ④, which was the best among the four unidirectional communication methods in the subjective evaluation. In the figure, we present an example in which three drink bottles are recognized. Methods ① through ⑥ verbally describe the recognition results following the policy of each method. Method ⑦ does not verbally describe the recognition results. Instead, it just plays a ping sound to let the user know the system recognized something. If the user wants to know what it is, the user can ask the system. Otherwise, the user can ignore it. Methods ⑤, ⑥, and ⑦ accept weak questions. The weak questions do not directly ask which product is cheapest, but by combining the answers of the questions, the user can obtain the information about the cheapest product. Methods ⑥ and ⑦ also accept a strong question that directly asks which product is cheapest. Note that in the experiment, we used the names of actual products instead of abstract expressions such as "Tea A."

munication methods, as shown in Fig. 1. Note that we call the methods that do not use interactive communication, such as the four methods introduced above, *unidirectional communication* by contrast.

In this paper, as a test bed of the PIA framework, we implemented a voice guidance system using a wearable camera with a wide field-of-view lens. It is equipped with the functions of the seven information summarization and interactive communication methods. We experimentally evaluated the seven methods with nine people with visual impairment.

2 Related Work

We have introduced an idea related to the PIA framework [8]. In it, we consider two questions: "What is the object?" and "Where is the object?" The answers to the questions are limited to either *known* or *unknown*. For simplicity, let us denote "what the object is" by *what* and "where the object is" by *where*. Then, on the basis of their answers, we categorize the situations in which people with visual impairment obtain the surrounding visual information into the following three types.

Category 1: *What* is unknown and *where* is known.
Category 2: *What* is known and *where* is unknown.
Category 3: *What* is unknown and *where* is unknown.

Category 1 is the case in which the user wants to know what the object is. That is, the user knows the location of the target object but does not know what it is. It is expected that this problem can be solved by using the existing apps and devices that require the user's active actions stated above. In Category 2, the user does not know where a specific object exists. A representative task for this category is finding something. As a possible solution for this problem, a method that uses an omnidirectional camera has been considered [8]. In Category 3, the recognition target is unknown, so the user has no clues about its location. One example of this situation occurs when the user comes across unexpected information while walking around town. Among three categories, Category 1 corresponds to the AIA framework, and Categories 2 and 3 correspond to the PIA framework.

In the rest of this section, we survey existing approaches to summarize information obtained from recognition techniques.

Pseudo-visual Attention Approach: Sighted people instantly evaluate the importance of information coming in through their eyes. If the visual attention mechanism of sighted people could be reproduced on a computer, only the important information would be selected and summarized. Some visual attention mechanisms have been modeled [3,15,18].

Information Recommendation Approach: Information recommendation provides information derived from the user's preferences [5,17]. However, it requires a large amount of data to estimate the user's preferences.

Information Theory Approach: Bracha et al. [2] proposed an information theoretic approach, which is an unsupervised method to determine which label should be given priority in order to increase the amount of information.

QA Approach: In the QA approach, the user should be able to directly ask for any information needed. As related work, visual QA (also known as VQA) returns the correct answer when presented with an image and a question related to the image [1,6,10].

3 User Study

We asked nine people with visual impairment (six with complete blindness and three with low vision) to perform a pseudo-shopping task. The detailed profiles

Table 1. Participants' demographic information.

ID	Age	Sex	Visual impairment	Onset age
P1	36	F	Totally blind	15
P2	30	M	Totally blind	0
P3	27	F	Totally blind	3
P4	28	F	Light perception	0
P5	35	M	Low vision (left: light perception, right:0.01)	20
P6	31	F	Low vision (left: 0, right: 0.01)	18
P7	27	M	Totally blind	5
P8	47	M	Totally blind	20
P9	57	M	Low vision (left&right: 0.02, visual field: 10°)	42

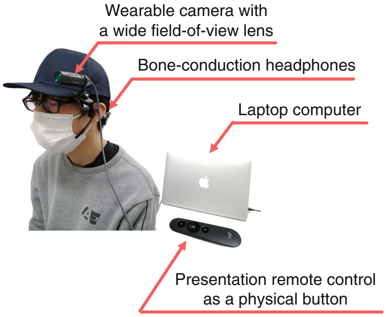
Wearable camera with a wide field-of-view lens
Bone-conduction headphones
Laptop computer
Presentation remote control as a physical button

Fig. 2. Voice guidance system.

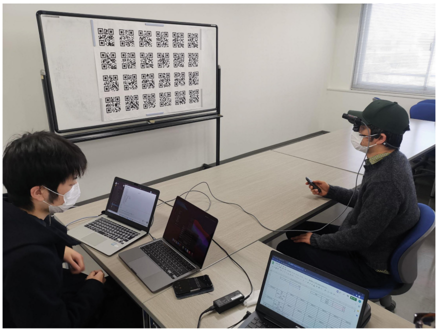

Fig. 3. Snapshot of the experiment.

of the participants are shown in Table 1. For quantitative evaluation, we asked them to perform the task of finding the cheapest product with the support of a voice guidance system, as shown in Fig. 2. The voice guidance system recognizes registered products every frame and basically reads aloud the names and prices of the recognized products. In the voice guidance system, we implemented seven information transmission methods that tell the user the recognition results differently, as shown in Fig. 1. Methods ⑤, ⑥, and ⑦ accept the following four weak questions: "What are the products that have been recognized?," "How much is (product name)?," "What are the categories that have been recognized?," and "What is the cheapest (category name)?" Methods ⑥ and ⑦ accept the following single strong question: "What is the cheapest product among all the recognized products?" The strong question is straightforward and easy to use for the user, but the system may not be ready to accept it; a strong question is specific and not frequently used in contrast to weak questions that are general and frequently used, which makes the cost and effort to prepare a strong question relatively high. Therefore, we also consider the case in which the user has to repeatedly ask weak questions until the desired information is obtained. This process places a relatively high cognitive load on the user.

A snapshot taken during the experiment is shown in Fig. 3. A participant wearing the voice guidance system sits on a chair in front of a table. A whiteboard is placed approximately 2 m away in front of the participant. An A0 paper on which 24 QR codes are printed is placed on the whiteboard. The QR codes are substitutes for products sold in a supermarket; when the voice guidance system recognizes a QR code, the system regards it as a registered product. We used QR codes to exclude as many effects caused by failure in recognition as possible. The voice guidance system runs on a laptop computer. Two experimenters, one of which appears on the left in the image, control the experiment. Although it is substantially different from shopping in actual daily life, we received feedback from some participants that the shopping in this experiment might be close to actual window shopping.

In the experiment, participants tested seven information transmission methods, referred to as methods ① to ⑦. Participants interacted with each information transmission method twice to buy the cheapest product of two different product types (i.e., a drink and snack). After the participants finished testing method ⑦, they were asked to test method ① again. This was done to evaluate the effect of the participants' familiarity; that is, how much the task completion time and the accuracy changed as the participants became familiar with the system. The second trial of method ① is referred to as method ①'.

3.1 Voice Guidance System

To evaluate information transmission methods, we implemented a voice guidance system using a wearable camera with a wide field-of-view lens, as shown in Fig. 2. This system was built using a combination of a wearable camera (Panasonic HX-A1H with a 150-degree field of view, 45 g), bone-conduction headphones (AfterShokz OPENCOMM, 23 g), and a presentation remote control (Logicool R500GR, 54 g) as a physical button. The wearable camera was connected by a cable to the laptop computer. The participant wore a cap, and the wearable camera was fixed with a mounting clip to the brim of the cap. The wearable camera reads the QR code that is associated with product information, and this information is described by voice through the bone-conduction headphones. Bone-conduction headphones were used for this experiment because it does not cover the ears and enables the participants to hear any instructions from the experimenters. When the participant felt uncomfortable or confused by listening to the large volume of product information, they could temporarily halt the verbal description at their own discretion using the physical button (presentation remote control). However, no participant used this function during the experiment. The verbal messages were spoken using Google Cloud Text-to-Speech. To recognize the QR codes, we used OpenCV and pyzbar, which are Python libraries.

Table 2. Subjective evaluation with a relative ranking. ">" indicates that the left is easier to use than the right, and "=" comparable. Hence, in general, the more to the left, the easier it is to use.

ID	Preference
P1	⑦ = ⑥ > ③ > ④ > ① > ② > ⑤
P2	⑦ = ⑥ > ⑤ > ④ > ③ > ② > ①
P3	⑦ > ⑥ > ④ > ⑤ > ③ > ② > ①
P4	⑦ > ⑥ > ⑤ = ④ > ③ > ② = ①
P5	⑦ > ⑥ > ⑤ > ④ > ② > ③ > ①
P6	⑦ = ⑥ > ⑤ > ③ > ④ > ② = ①
P7	⑦ = ⑥ > ⑤ > ③ > ④ = ② > ①
P8	⑦ > ⑥ > ⑤ > ④ > ③ > ② > ①
P9	⑦ > ⑥ > ④ > ③ > ① > ② > ⑤

3.2 Results and Discussion

The information transmission methods were subjectively and objectively evaluated.

Subjective Evaluation. Subjective evaluation of the information transmission methods was performed based on a relative ranking with comments. As mentioned above, seven information transmission methods were tested: methods ① to ⑦. After testing each method, an experimenter asked the participant about the relative rank of the method regarding ease of use and to comment on the method. Accumulating the relative ranks of the seven information transmission methods, we obtained the participants' preferences, as shown in Table 2.

Let us first focus on unidirectional communication methods (i.e., methods ① through ④). The table shows that almost all participants stated that methods ③ and ④ were easy to use. Both methods use semantic summarization, and therefore this proves the effectiveness of semantic summarization. In addition, comparing methods ① with ② and methods ③ with ④, methods ② and ④ were found to be preferred by most participants. Methods ② and ④ use temporal summarization, and therefore this proves the effectiveness of temporal summarization.

Next, we consider all methods. Comparing unidirectional communication methods and interactive communication methods, the table shows that interactive communication methods (i.e., methods ⑤ through ⑦) were preferred over unidirectional communication methods (i.e., methods ① through ④). All participants stated that method ⑦ was the best and method ⑥ was the second best, both of which allow the use of a strong question. By contrast, six participants ranked method ⑤ in third place, whereas three ranked lower (i.e., one ranked it in fourth place, and two ranked it last). This is because method ⑤ only allows

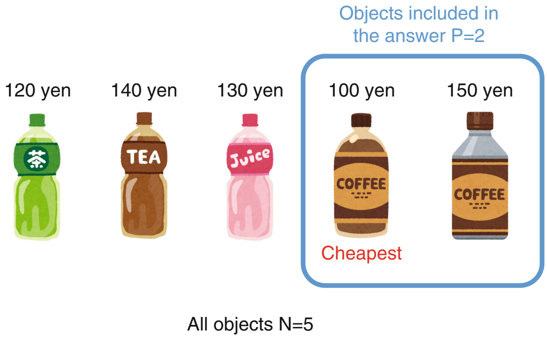

Fig. 4. Example calculation of the score defined by Eq. (1).

the use of weak questions. This implies that combining weak questions is not always comfortable for the users.

Objective Evaluation. The objective evaluation of the information transmission methods was performed based on the average time to complete the task and its accuracy, represented by a score.

The time to complete the task is defined as the time duration between the first recognition began on the voice guidance system and the participants raised their hand to state the name and price of the cheapest product. However, the experiment was terminated when the time exceeded 180 s, and the participants were requested to give their answer. The accuracy was measured by the score, which was defined in the range of 0% (the lowest) to 100% (the highest). Score Y is defined by

$$Y = \frac{1}{2}\frac{\log N - \log P}{\log N}X_{\mathrm{name}} + \frac{1}{2}X_{\mathrm{price}}, \tag{1}$$

where X_{name} and X_{price} respectively represent whether the answers for product name and price are correct; they are 1 if the answer is correct and 0 if it is incorrect. Furthermore, N is the number of all products and P is the number of the products not eliminated by the answers (see also Fig. 4), as described in detail below. The participants often described the product ambiguously, for instance, "It was a tea-type drink for 120 yen." This ambiguity had two causes: 1) the existence of long or similar product names; and 2) the participants tended to forget the product names when they focused on the prices. Thus, judging an answer as correct or incorrect hinders the meaningful evaluation of the accuracy. Therefore, we take into account the intermediate states in the manner of information theory. For example, as shown in Fig. 4, we consider the case in which the participant answered, "Some coffee is the cheapest for 100 yen." In this case, from the answer, we can reduce the candidate products from all $N = 5$ objects to $P = 2$ objects, and the cheapest product is included in the candidate products. Thus, $X_{\mathrm{name}} = 1$. In addition, the participant's answer for the cheapest

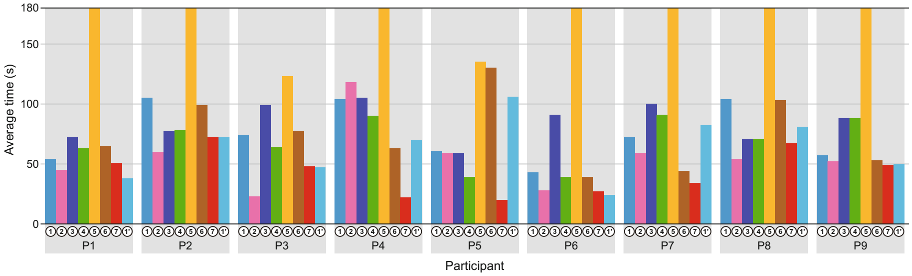

Fig. 5. Average time to complete the task for each participant.

price is correct, thus $X_{\text{price}} = 1$. Following the idea of information theory, the answer reduced the ambiguity of $N = 5$ products to the ambiguity of $P = 2$ products. Therefore, the information provided by the answer is $\log N - \log P$. To normalize this score so that it ranges between 0 and 1, we divide it by its maximum value, $\log N$, to obtain the first term of (Eq. 1). As a consequence, the score is $Y = 0.78$ in this case.

Regarding the average time to complete the task, we found that it depends more on the participants than the information transmission methods. Therefore, we visualized the average time to complete the task for each participant in Fig. 5. Focusing on unidirectional communication methods, Participants P1, P5, P6, and P9, whose vision became impaired to its current level relatively recently, required around 60 s to complete the task, which was less than the times of the other five participants. Focusing on interactive communication methods, all participants except for P3 and P5 spent 180 s for method ⑤. When using that method, to obtain the cheapest product, the participant repeatedly needs to combine two questions: “What are the categories that have been recognized?” and “What is the cheapest (category name)?” This process places a relatively high cognitive load on the user because the user needs to remember the category list as well as the name and price of the cheapest product thus far. By contrast, methods ⑥ and ⑦, which allow the use of a strong question, required much less time. Owing to the existence of the strong question, the participants could easily identify the cheapest product using the strong question. This implies the power of interactive communication methods. Comparing methods ① and ①' (the second trial of method ①), all participants except for P5 and P7 reduced the time in the latter trial. On average, it was reduced by 13%. This indicates that the participants quickly became used to the experiments. However, even taking this effect into account, the consideration mentioned above is not affected. We predicted that the participant with congenital blindness might be more tolerant to verbal information acquisition, but we did not confirm such a correlation. We received feedback that participant P6, who had low vision, had seen the product used in the experiment before, so that it was much easier to imagine the product and obtain the information.

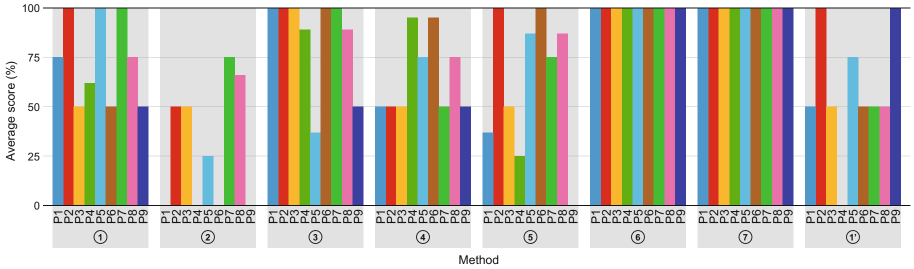

Fig. 6. Average score for each information transmission method.

Figure 6 visualizes the average score for each information transmission method. Let us focus on unidirectional communication methods first. We obtained different results depending whether temporal summarization was used. The scores of methods ② and ④, which use temporal summarization and a verbal description only once were low, whereas those of methods ① and ③, which do not use temporal summarization and verbally describe the results repeatedly were high. It is believed that this score difference is explained by the fact that the participants were not allowed to reconfirm the product name when they failed to catch it or forgot it. A notable result was obtained for method ②, for which the scores of four participants were 0. This seems to be a typical disadvantage of temporal summarization. Next, let us focus on interactive communication methods. Methods ⑥ and ⑦, which allow the use of a strong question, achieved 100% for all participants. This implies the ease of use of the strong question. In contrast, the scores of method ⑤, which allow the use of only weak questions, were worse than those of methods ⑥ and ⑦. A significant difference in the scores depending on participants arose because of the termination of the experiment at the upper time limit; when the time exceeded 180 s, participants were requested to give an answer even if they had no idea which answer was correct.

4 Conclusions

With regard to using smartphone apps and assistive devices supporting people with visual impairment to obtain visual information from the surroundings, this paper pointed out two problems that have not attracted much attention. The first problem is that the user is required to actively take a picture as an implicit precondition. However, this is not always possible for people with visual impairment. Therefore, we proposed a new framework called *passive information acquisition (PIA)* that is independent of the active actions performed by the users. The proposed PIA framework can be actualized by constantly recording the user's surroundings and recognizing all the images. A downside of this framework is that it can obtain a huge amount of surrounding visual information, which may overwhelm the user. Thus, information summarization is crucial. Another problem is that the user must constantly listen to the voice describing aloud the recognition results when using assistive apps and devices. Compared with sighted people, who can focus on the information of interest without evaluating everything in their sight, people with visual impairment must use methods that are far less efficient. A possible solution to this problem is to introduce *interactive communication* using a question answering (QA) system, where the user requests information from the recognition system, and then the recognition system tells the requested information to the user.

We conducted an experiment with nine people with visual impairment and asked them to perform a pseudo-shopping scenario. In the experiment, we examined four unidirectional communication methods that were designed to examine all four combinations of with/without temporal/semantic summarization. These methods were subjectively evaluated with relative rankings and objectively evaluated based on the task completion time and accuracy. As a result, the method with both temporal and semantic summarization was the best in the subjective evaluation. Although this method was not the best in the objective evaluation, the effectiveness of temporal summarization was confirmed.

We also examined three interactive communication methods. Comparing the interactive communication methods with the unidirectional communication methods, we found that the methods that allow a strong question to be used were the best in the subjective evaluation. These methods were also the best with respect to accuracy in the objective evaluation. However, the method that allows only weak questions was not always better than the best unidirectional communication method both in the subjective and objective evaluations.

Acknowledgments. This work was supported by JSPS Kakenhi Grant Number 20H04212.

References

1. Antol, S., et al.: VQA: visual question answering. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2425–2433 (2015)
2. Bracha, L., Chechik, G.: Informative object annotations: tell me something i don't know. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
3. Broadbent, D.E.: Perception and Communication. Elsevier, Amsterdam (2013)
4. Chiu, T.Y., Zhao, Y., Gurari, D.: Assessing image quality issues for real-world problems. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3643–3653 (2020). https://doi.org/10.1109/CVPR42600.2020.00370
5. Davidson, J., et al.: The youtube video recommendation system. In: Proceedings of the Fourth ACM Conference on Recommender Systems, pp. 293–296 (2010). https://doi.org/10.1145/1864708.1864770
6. Gurari, D., et al.: VizWiz grand challenge: answering visual questions from blind people. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)

7. Iwamura, M., Hirabayashi, N., Cheng, Z., Minatani, K., Kise, K.: VisPhoto: photography for people with visual impairment as post-production of omni-directional camera image. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–9 (2020). https://doi.org/10.1145/3334480.3382983
8. Iwamura, M., Inoue, Y., Minatani, K., Kise, K.: Suitable camera and rotation navigation for people with visual impairment on looking for something using object detection technique. In: Proceedings of the 17th International Conference on Computers Helping People with Special Needs (ICCHP 2020) (2020). https://doi.org/10.1007/978-3-030-58796-3_57
9. Jayant, C., Ji, H., White, S., Bigham, J.P.: Supporting blind photography. In: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 203–210 (2011). https://doi.org/10.1145/2049536.2049573
10. Jiang, H., Misra, I., Rohrbach, M., Learned-Miller, E., Chen, X.: In defense of grid features for visual question answering. In: Proceedings of the CVPR (2020)
11. Kacorri, H., Kitani, K.M., Bigham, J.P., Asakawa, C.: People with visual impairment training personal object recognizers: feasibility and challenges. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM Press (2017). https://doi.org/10.1145/3025453.3025899
12. Kayukawa, S., et al.: BBeep: a sonic collision avoidance system for blind travellers and nearby pedestrians. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–12 (2019). https://doi.org/10.1145/3290605.3300282
13. Kayukawa, S., Ishihara, T., Takagi, H., Morishima, S., Asakawa, C.: BlindPilot: a robotic local navigation system that leads blind people to a landmark object. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–9 (2020). https://doi.org/10.1145/3334480.3382925
14. Kayukawa, S., Takagi, H., Guerreiro, J.A., Morishima, S., Asakawa, C.: Smartphone-based assistance for blind people to stand in lines. In: Proceedings of the CHI Extended Abstracts, pp. 1–8 (2020). https://doi.org/10.1145/3334480.3382954
15. Lavie, N., Tsal, Y.: Perceptual load as a major determinant of the locus of selection in visual attention. Percept. Psychophys. **56**(2), 183–197 (1994)
16. Lee, K., Hong, J., Pimento, S., Jarjue, E., Kacorri, H.: Revisiting blind photography in the context of teachable object recognizers. In: Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 83–95 (2019). https://doi.org/10.1145/3308561.3353799
17. Thorat, P.B., Goudar, R., Barve, S.: Survey on collaborative filtering, content-based filtering and hybrid recommendation system. Int. J. Comput. Appl. **110**(4), 31–36 (2015)
18. Treisman, A.M.: Contextual cues in selective listening. Q. J. Exp. Psychol. **12**(4), 242–248 (1960)
19. Vázquez, M., Steinfeld, A.: Helping visually impaired users properly aim a camera. In: Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 95–102 (2012). https://doi.org/10.1145/2384916.2384934