



AR Assistive System in Domestic Environment Using HMDs: Comparing Visual and Aural Instructions

Shuang He^{1,2,3(✉)}, Yanhong Jia^{2,3}, Zhe Sun³, Chenxin Yu³, Xin Yi¹,
Yuanchun Shi¹, and Yingqing Xu^{2,3}

¹ Department of Computer Science and Technology,
Tsinghua University, Beijing, China
heshuang.design@foxmail.com,
{yixin, shiyu}@mail.tsinghua.edu.cn

² Department of Information Art and Design,
Tsinghua University, Beijing, China
jyhl8@mails.tsinghua.edu.cn,
yqxu@mail.tsinghua.edu.cn

³ The Future Laboratory, Tsinghua University, Beijing 100084, China
{sunz15, yu-cx16}@mails.tsinghua.edu.cn

Abstract. Household appliances are becoming more varied. In daily life, people usually refer to printed documents while they learn to use different devices. However, augmented reality (AR) assistive systems providing visual and aural instructions have been proposed as an alternative solution. In this work, we evaluated users' performance of instruction understanding in four different ways: (1) Baseline paper instructions, (2) Visual instructions based on head mounted displays (HMDs), (3) Visual instructions based on computer monitor, (4) Aural instructions. In a Wizard of Oz study, we found that, for the task of making espresso coffee, the helpfulness of visual and aural instructions depends on task complexity. Providing visual instructions is a better way of showing operation details, while aural instructions are suitable for presenting intention of operation. With the same visual instructions on displays, due to the limitation of hardware, the HMD-users complete the task in the longest duration and bear the heaviest perceived cognitive load.

Keywords: Augmented reality · Multimodal feedback · Assistive system · Instructions · Head mounted displays

1 Introduction

In daily life, people usually refer to printed documents while they complete tasks by using household appliance, for instance, using a coffee machine to make espresso coffee according to the manual. During the course of operation, they may skip steps, misunderstand instructions, or unable to correct mistakes in time. In part, these problems are caused by the readability of the documents. On the other hand, the gap between the paper instructions and the actual context reduces the effectiveness of

operation. We want to build an eyes-free, hands-free and non-distractive external learning environment for users. Therefore, AR assistive system providing visual and aural instructions based on head mounted displays seemed to be an ideal solution.

In this paper, we conducted a heuristic evaluation and a think aloud protocol user study based on paper manual to establish user's mental model as well as instruction design principles. We redesign the paper document according to the experimental task and develop three prototypes of assistive system, one with HMDs, one with computer monitor and one with conversational voice system. All three of the prototypes offering support to users with the same quantity of information. Through a Wizard of Oz user study with 20 participants, we compared users' performance for instruction understanding in visual and aural modality.

We have made three main contributions in this paper: (1) a diagram applied to design multimodality AR assistive system: When system providing visual or aural detailed description for each step, a simple and clear statement should contain three key information points, target objects, relative position between target objects and actions. (2) a design principle based on task complexity: For novices, when they use home appliance to perform some easy functions, there is no significant difference in helpfulness between visual and aural instructions. However, when a step contains several complex operations, visual instructions provide more helpfulness. Comparing to graphic user interface, a well-designed voice user interface is more flexible. Voice user interface is suitable for presenting intention of operation and users demand supplementary information from conversational virtual assistant expecting for better interactive experience. (3) a method applied to analyze users' decision-making process on visual and aural instructions, helping researchers to design multimodal AR assistive system.

The result of user study shows that, due to the limitation of the hardware, AR assistive system based on HMDs have theoretic feasibility on current stage. To reach the ideal state, it requires better recognition and display technology. Conversational assistive system could effectively reduce perceived cognitive loads. But when users are unfamiliar with the devices, descriptiveness of aural instructions is limited. With the accumulation of hands-on experience, there is less demand for detailed description of each step and users tend to seek support from the system for specific problems rather than being instructed.

2 Related Work

In recent years, research on AR assistive system involves varies domains, such as hospital setting [1, 2], remote work support [3], industrial manufacturing [4], education applications [5], etc.

In earlier research, AR multimodality assistive system in human settlement is usually built with in-situ projection. CounterActive is a cooking aid system used in the kitchen guiding the users with projection and aural instructions [6]. In the work of Ayaka Sato et al., the MimiCook system uses image recognition technologies to analyze user activity. With in-situ projection, Mimicook displays menu as well as supplementary information to improve users' task efficiency [7]. Yu Suzuki and

Shunsuke Morioka et al. developed a cooking support system for novices, including a conversational robot assistant “Phyno”. Phyno can interact with the user via voice and gestures [8]. Meanwhile, there are some mixed modality conversational virtual assistants coming on the market, such as Siri, Amazon Echo. The aim of the research above are not to report interactive efficiency or perceived cognitive loads, but rather focus on system design, implementation, user experience and acceptance of introducing AR assistive system in domestic environment.

With the development of AR technology, the research reports task completion time, error frequency and interactive efficiency of AR assistive system in specific scenario is gradually unfolding. Markus Funk et al., compared assembly instructions based on HMDs, tablets, in-situ projections and baseline paper documents. The results show that for assembling tasks, completion time is significantly longer using HMDs. HMD-users make more error and have more perceived cognitive load [9, 10]. Markus Funk’s research also involves the comparison of different visualizations, such as pictures, videos, 3D models and contour. They report that contour visualization is significantly better in perceived mental load and performance of the impaired participants [2].

For multi-modality feedbacks, Marina Cidota et al. compared audio and visual notifications in remote workspace collaboration. After analyzing the case of placing virtual objects in the shared workspace, they find that visual notifications are preferred over audio or no notifications independent from the level of difficulty of the task [11]. Youngsun Kim et al. presented an AR-based tele-coaching system for fast pace task applied to the game of tennis. They evaluate the instantaneous response rate of visual, voice, and multimodal augmented instructions. Sound show the worst performance in terms of the responsive time. AR is the most useful in stringent temporal conditions. Multimodal feedback seems to caused distraction to users [12].

Overall, previous work investigated the acceptance of AR assistive system in domestic environment. However, a comprehensive study comparing visual and aural instructions using HMDs in domestic environment has not been done yet. In this paper, we will compare baseline paper instructions, aural instructions to visual instructions based on HMDs and computer monitor. Further, we will evaluate user performance following long instructions in complex tasks in human settlement.

3 Initial Study

3.1 Evaluating Paper Document

We conducted a heuristic evaluation of paper documents in accordance with Nielson’s usability principles [13, 14]. Five researchers with the background in human-computer interaction participated in the test. They evaluated the usability of original paper instruction of De’Longhi ECO 310 Icona Espresso Manual Machine (Chinese) [15]. Then, we launched a think aloud protocol test [16] involving 3 participants. They were asked to make espresso coffee according to the manual.

The results of the heuristic evaluation and the think aloud protocol test indicate that, for the task of making espresso coffee, the problems on usability and readability in original manual are found as follows: (1) The operative description in paper manual cannot be fully matched to actual operating procedure. (2) The paper documents

include terms that are not familiar to novices (i.e. filter-holder, extraction, etc.). (3) Semantic ambiguity in descriptive text, icons, and diagrams. (4) Information redundant: precautions, explanatory and descriptive information have been cropped up in operating procedures. (5) User need to follow the serial number among instruction text to look up for diagrams, which may easily lead to misreading or skipping steps.

In general, we need to select a task for experimentation, clear the procedures and simplify the instructions. The usability and readability problems in paper manual should be revised. Besides, related diagrams and text descriptions for each step should be form a one-to-one correspondence and displayed on same page.

3.2 Case Study of Making Espresso Coffee

We clarified that the aim of experimental task was to “making a cup of espresso coffee with a manual coffee machine”. Operating procedure was simplified into 9 steps with intention and specific operative descriptions (see Table 1).

The task of “making espresso coffee” was chosen because: (1) It involves the primary operation methods for the coffee machine, helping users to understand the operating principle. (2) It includes numerous operative steps and appliance, but makes limited use of ingredients. Therefore, users’ prior experience will not affect their performance. (3) Concerning task complexity, operation has different levels of difficulty. Therefore, the task is considered appropriately to be simulated in the laboratory environment.

The instructions were grouped in to two categories: (1) Simple instruction: it describes how to interact with a single object. The spatial position or state of the object changes such as “Position the cup under the filter holder spouts”. (2) Complex instruction: it describes how to interact with more than two objects. It involves the changes of relative spatial position between objects and multiple physical feedback, such as “Attach filter holder into boiler outlet. Turn right to lock into position.”

Table 1. How to prepare espresso using ground coffee.

Step	Intention	Description	Complexity
1	Turn the machine on	Press the ON button	Simple
2	Install steel filter	Install steel filter cup into the filter holder	Complex
3	Filling coffee	Fill the filter with a level measuring scoop of coffee	Complex
4	Tamp the coffee	Press the coffee lightly using the coffee tamper	Complex
5	Attach filter holder	Attach filter holder in place into boiler outlet. Turn right to lock into position	Complex
6	Place coffee cup	Position the cup under the filter holder spouts	Simple
7	Extraction	Press the coffee button. (It is recommended not to run the coffee for more than 45 s). Press the same button again	Simple
8	Cleaning	Remove the cup. Turn the grip from right to left to release the filter holder	Simple
9	Turn the machine off	Press the OFF button	Simple

3.3 Tutorial Design Paradigm

In the analysis of procedures and instructions, we generalized that when assistive system providing visual or aural instruction for each step, a simple and clear statement should include three key information points: (1) Objects, what will be used in operating process; (2) Relative position between objects, where the objects should be placed; (3) Actions, how to manipulate. For example, “Attach filter holder in place into boiler outlet”. The objects involved are “filter holder” and “boiler outlet”, the action is “attach into”, and the relative position is from separating to assembling.

In order to avoid affecting users’ performance, instructions in paper document, visualization and auditory need to convey the same amount of information. In visual instructions, an object is marked with a circle. We draw lines between objects, which started with a dot and ended with an arrow, to indicate relative position and orientation. Action is described by a dynamic arrow. Orientation and acceleration of the arrow indicates how much force should be performed. Intention of each step is displayed in the lower left corner in text. In conversational voice system, aural instructions are the precise description of dynamic graphic illustrations (see Fig. 1).



Fig. 1. Example of instructions. (1) Paper manual, (2) Visual instruction, (3) Aural instruction.

4 System

For exploring visual and aural instructions in human settlement, we simulate a domestic environment in independent space in our lab. In the following, we introduce the system prototype of visualizations and auditory.



Fig. 2. AR assistive system. (1) Live stream, (2) System being used, (3) Instruction on HoloLens.

4.1 Visualizations

Visual instructions were designed according to the diagram described in Sect. 3.3. Previous works suggested that compared to video and pictorial visualization, contour instructions resulted in fewer errors and better performance. In our previous study, we tried to make contour instructions on HMDs. We used HoloLens Vuforia SDK for image recognition and built dynamic contour with Unity 3D. Because the speed of image recognition and tracing was rather slow, we decided to use video instructions instead. We photographed the standard task procedures and overlaid dynamic contour instructions on the image to simulate contour-overlapping.

Visual Instruction Based on HMDs. For HMDs implementation, the instruction system was developed with Unity 3D and C#, then implemented on HoloLens. In this system, a serial of silent video clips was provided along with the serial number of each step, the intention of the operation and minimum additional text explanation if needed (only a few steps have such additional explanation, i.e. It is recommended not to run the coffee for more than 45 s). Users could switch to the next or the previous step by interacting with buttons on the right or left side of the video clips using gestures (see Fig. 2).

Visual Instruction Based on Screen. Considering that HMDs themselves might have influence upon the study, we implemented visual instructions not only in HMDs but also on a computer monitor (27-in. screen). For visual instructions on screen, video clips are displayed on a computer monitor orderly. Each clip shows the instruction of one step out of nine (see Table 1) with the serial number and the intention of the step. Once the user complete current step, the video of next step would be played on the screen.

4.2 Aural Instructions

The conversational voice system prototype was designed according to task procedure. Conversation samples were listed and dialogue process was analyzed based on task flow. According to task procedure (see Table 1), we referred to existing voice-based interface design paradigm and developed a preset list of aural instructions and responses. Experimenter (the wizard) could select computer synthesized audio clip from the response list. We transferred text to voice instructions by Responsive Voice API¹.

After observing users' interaction in pre-test with the voice-based instruction system, we classified user behavior into five categories: Explicit next, Request, Implicit next, operation errors or timeout and undefined response. Different responses were given to the user in a Wizard of Oz study according to this classification (see Table 2).

¹ <https://responsivevoice.org/api/>.

Table 2. Example of user behavior and system response

User behavior	Response
Explicit Next: “ <i>Okay</i> ”, “ <i>Next?</i> ”	Play the instruction of next step
Request: “ <i>What is...?</i> ” “ <i>Say it again?</i> ”	Repeat the instruction of current step
Implicit next (complete and wait quietly)	Play the instruction of next step
Errors or timeout	Repeat the instruction of current step
Undefined response	“ <i>Sorry, I am listening</i> ”

5 Evaluation

In this work, we developed a high-fidelity Wizard of Oz simulation to evaluate user performance. We describe the protocol and apparatus below.

5.1 Procedure

We invited 20 participants to engage in our user study. The participants were aged between 20 to 33 (Avg. = 25, SD = 4.8); 9 participants were male, 11 were female. Participants were divided into four groups. Each group was asked to take one of the four evaluations: (1) Paper instructions, (2) Visual instructions based on HMDs, (3) Visual instructions based on computer monitor, (4) Aural instructions. All participants have no experience of using a manual coffee machine. Out of the 5 participants that attended HMDs evaluation, 2 reported having used HMDs within half a year, 2 reported earlier, 1 never; while among the 5 that attended aural instructions evaluation, 1 reported using a conversational assistant usually, 2 reported having used an intelligent voice system within half a year.

The experiment took place in an independent space at our research facility. Participants were briefed upon arrival, and were given 5 min to read a paper introduction about the components of the coffee machine that would be used in the following experiment. Participants of visual instruction group were guided by the same serial of looping video clips. Switching from one step to another could be done by gestures. An additional operation was provided for HMDs users to replace the video clips to wherever they prefer. Participants of conversational assistive system were informed that they could communicate with the system, either asking for step changing or instruction repeating. All participant actions were audio and video recorded.

During the experiments, lab assistants wouldn't be involved unless necessary (e.g. If users tried to do something that would probably make themselves in danger, assistants would intervene). Participants were asked to complete System Usability Scale (SUS) questionnaires [17], the NASA Task Load Index questionnaire (NASA-TLX) [18] and a semi-structured interview. The post-interview mainly focuses on: (1) Overall feelings about the system and its advantages and disadvantages; (2) Causes of errors or confusion during the experiment; (3) Open interview according to results from SUS and NASA-TLX, where participants were asked to give suggestions to better user experience.

5.2 General Impressions

All of the 20 participants completed the task. Visual instructions on screen received the highest SUS average score of 70.83 (SD = 13.91), followed by aural instructions with an average score of 69.5 (SD = 9.82) and paper instructions with an average score of 67.5 (SD = 25.74). AR assistive system based on HMDs was the least favorable, which got the lowest average score of 66.5 (SD = 9.62). All of the four prototypes were performing at the acceptable threshold level.

5.3 Task Completion Times

We accumulated users’ response time and the procedure duration at each step. For the length of instructions varied a lot, we used the “point-in-time” when users started to act as the starting point for timing. When a step was completed, we stopped timing.

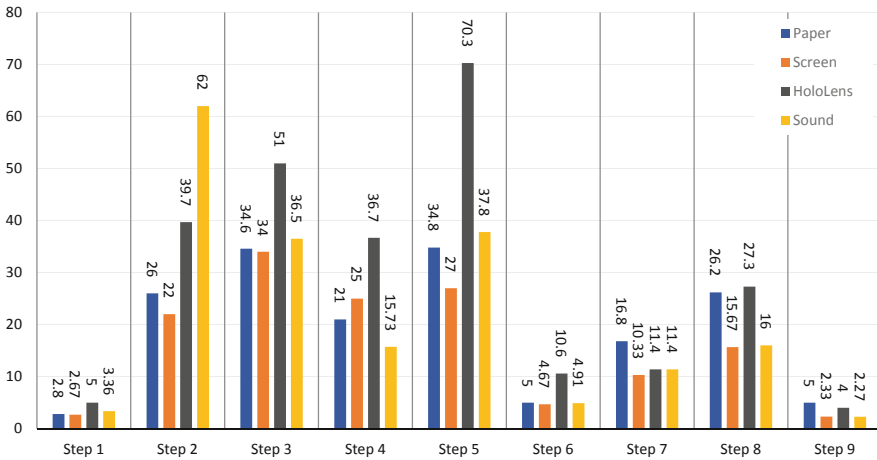


Fig. 3. Overview about the results of step completion times. Comparing (1) Paper instructions, (2) Visual instructions based on HMDs, (3) Visual instructions based on screen, (4) Aural instructions.

Participants referred to visual instructions on screen taking an average of 143.7 s (SD = 52.7) in finishing the task, which was the fastest, followed by the group using paper document with an average of 172.2 s (SD = 19.6), aural instructions with an average of 190.0 s (SD = 58.6) and visual instruction on HMDs with an average of 256 s (SD = 58.6).

For an overview of each step completion times (see Fig. 3.), participants who was supported by HMDs possessed longer procedure duration than users who was instructed by screen. When users performed some easy functions (i.e. Step1, 6, 7, 8, 9), the there was no significant difference in helpfulness between visual instructions on screen and aural instructions. However, when a step contains several complex operations (i.e. Step2, 3, 5), visual instructions seemed to improving operative efficiency.

5.4 Errors

On the whole, participants made least mistake using visual instructions on screen with an average error rate of 0.07 (SD = 0.06), followed by aural instructions with an average error rate of 0.08 (SD = 0.1), paper instructions with an average error rate of 0.18 (SD = 0.11) and visual instructions on HMDs with an average error rate of 0.36 (SD = 0.21).

We further analyzed the errors participants made while completing the task in each step. Participants supported by paper manual made errors in Step2 (error rate = 0.4), Step5 (error rate = 0.6), Step7, Step8 and Step9 (error rate = 0.2). Except for Step8, participants using the HMDs instructions made mistakes in all the other operations. Error rate for Step2 and Step5 was as high as 0.8. Users supported by visual instructions on computer monitor made mistakes mainly at Step2 (error rate = 0.4) and Step4 (error rate = 0.2). Errors occurred to conversational system users at Step2 (error rate = 0.4), Step3 (error rate = 0.2) and Step8 (error rate = 0.2).

5.5 Cognitive Load

The participants using aural instructions had the least perceived cognitive load with an average score of 6.73 (SD = 2.92). HMD-users reported the highest perceived cognitive load with an average TLX score at 10.01 (SD = 0.98). Meanwhile, visual instructions on screen caused the perceived cognitive load with a score of 8.27 (SD = 3.83). Paper instructions got the average score of 9.6 (SD = 5.48). We also analyzed the perceived cognitive load on six dimensions with each instruction techniques (see Fig. 4).

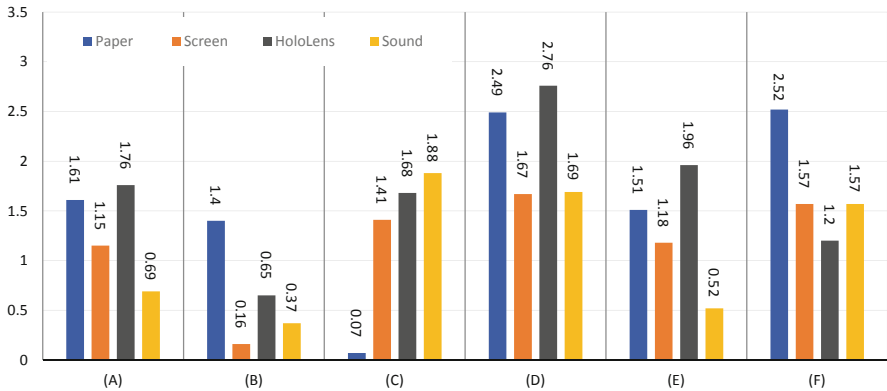


Fig. 4. Overview about the results of NASA-TLX. (A) Mental demand, (B) Physical demand, (C) Temporal demand, (D) Performance, (E) Effort, (F) Frustration.

Mental Demand. Conversational system prototype perceived the lowest mental demand score on average of 0.69. AR assistive system on HoloLens led to the highest mental demand of average score of 1.76, followed by paper instructions with average score of 1.61 and visual instruction on screen with average score of 1.15.

Physical Demand. The participants perceived the paper manual reported the highest physical demanding feedback with an average score of 1.4. Meanwhile, three other group's report on physical demanding was rather low.

Temporal Demand. Paper instructions led to the lowest temporal demand with an average score of 0.07. The other three instruction techniques were considered to a higher temporal demand.

Performance. Participants perceived their performance best using the HMDs with an average score of 2.76, followed by paper instructions with 2.49. Participants supporting by visual instruction on screen and aural instructions perceived their performance less successful with an average score of 1.67 and 1.69.

Effort. The participants perceived the lowest effort using conversational system prototype with an average score of 0.52. The group using HMDs perceived their effort the highest with an average score of 1.96.

Frustration. Surprisingly, participants using HMDs perceived least frustration with an average score of 1.2, followed by visual instruction on computer monitor and conversational system both with 1.67. Participants using paper manual perceived the highest frustration with an average score of 2.52.

5.6 Qualitative Results

Additionally, we observed the process participants interacted with the systems. Besides the quantitative results we also collected qualitative result from post interviews.

Paper Instructions. Participants who used paper manuals casually browsed through the instructions before the experiment. When they encountered with problems, they would look up the context thoroughly. Most of the participants thought that the paper manual was simple and easy to understand. *"I think this manual is better than most manuals I have ever used. For there's not much redundant information and the description is quite clear."*

Visual Instructions on Screens. Supported by visual instructions on screen, participants completed the task without much effort. The main complaint was about the motion graphic instructions. *"I didn't notice the text in the lower left corner. I was busy watching the dynamic image. Coffee tamper merged into the background, so I had some difficulties in recognize it. In addition, I hope the it could show me how much force I should use while tamping."*

Visual Instructions on HMDs. Most of the inconveniences HMDs users encountered were caused by the hardware. Due to the insensitivity of gesture recognition, participants needed to click multiple times to switch step. Furthermore, even though we had provided the function of dragging and dropping the video to a better viewing position, few participants adjusted the viewing distance. *"I was gazing at the video on HoloLens. Sometimes I lose the video instruction in my view."* *"I am short sighted. I think the line of contour better be thicker."*

Aural Instructions. Without any visual instructions, participants who was supported by sound focused on the items to be handled in the task. If they were familiar with the items mentioned in the instructions, they would operate while listening to supplementary description. Otherwise, they'd ask system to repeat. Participants found that it was interesting being instructed by a conversational virtual assistant. *"I think it would be better if the system was able to play some music for me while I was waiting for my coffee."* *"I'd like to know how to discern good coffee from bad. But she didn't replied."*

6 Discussion

6.1 Comparison Among Modalities

Paper Document. Participants are able to go back and forward freely through paper manual, which is natural but time consuming. Matching images and text explanations are shown on paper instructions at the same time, giving more information than the other three prototypes. As reading paper instructions is the most familiar way to absorb information of how to use a new machine, users perceived temporal demand turns out to be considerably low. For the same reason, people will get frustrated easily once failed to complete tasks.

HMDs. We suggest that the low performance of HMDs was mainly caused by the hardware, instead of the design of contour visual instructions. Although the error rate of HMDs users is the highest among the four prototypes, the TLX frustration score of which turns out to be the lowest, implying that users are willingly to explore this new device. In further study, we would introduce smartphones as another carrier and launch more comparing experiments.

Auditory vs. Visualization. Compared to visual instructions, results on aural instructions are better than expected. While user listening to aural instruction, they focus on comprehension without any distractions. We point out that appliances being used have significant influence on the users' cognition. In cases where the appliances are familiar to the users, it is easy for them to understand the intention of certain operation. It explains the result that during some steps, users using voice system performed better than those who was visually supported. In such situation, aural system shares visual perceived cognitive load. While in cases where the appliances are unfamiliar to the users (i.e. steel filter, filter holder, coffee tamper, etc.), aural system gave out worse performance than visual system. In such situation, visual system provides more effective guidance than aural system does.

6.2 Design Method for AR Tutorial

According to our study, we propose an ideal design method for multi-modal AR tutorial: (1) Define learning purpose and contents; (2) According to the contents and the design paradigm in Sect. 3.3, list steps of operations and instructions. Description of text and diagrams should be as simple and basic as possible; (3) Extract objects to be handled in each step and launch semantic tests upon these objects, to build users'

mental modal; (4) Provide instruction from two aspects: On one hand, present the intention of each operation and the descriptive explanation of unfamiliar objects with voice. On the other hand, display spatial relationships between objects and tips for the operation with image. As for those complex operations, we recommend combining of visual and aural instructions.

7 Conclusion

In this paper, we explored the tutorial design method for AR assistive system applied to home appliances and evaluated different systems for providing instructions in domestic environment. We compared visual instructions on HMDs and screen to aural instructions and baseline paper instructions. Our result show that, well designed and displayed visual instructions provide helpful information. Aural instructions share visual perceived cognitive task load. Especially when users are familiar with the appliances, conversational assistive system is appropriate for building an eyes-free, hands-free and non-distractive external learning environment.

Although HMD-instructions have problems being perceived the highest cognitive loads and the longest procedure duration, with the improvement of image recognition accuracy and the acceleration of tracing speed, the results will be different. In future work, we want to investigate the effects on smartphone and portable devices by constructing multimodality assistive system, and explore users' response upon long instructions in both visual and aural modality.

References

1. Susanna, N., Björn, J.: Fun and usable: augmented reality instructions in a hospital setting. In: *OzCHI 2007 Proceedings*, pp. 123–130. ACM Press (2007)
2. Susanna, N., Björn, J.: Acceptance of augmented reality instructions in a real work setting. In: *CHI 2008, Extended Abstracts*, pp. 2025–2032. ACM Press (2008)
3. Naoya, M., TaRemote, K.: Instruction system using see-through augmented reality. In: *Proceeding SA 2015, 2015 Posters (2015)*. Article No.31
4. Markus, F., Andreas, B.: Comparing projected in-situ feedback at the manual assembly workplace with impaired workers. In: *PETRA 2015*. ACM Press (2016). Article No.1
5. Yingxue, Z., Siqi, L.: ChinAR: facilitating Chinese Guqin learning through interactive projected augmentation. In: *Chinese CHI 2015 Proceedings of the Third International Symposium of Chinese CHI*, pp. 23–31 (2015)
6. Wendy, J., Rebecca, H.: CounterActive: an interactive cookbook for the kitchen counter. In: *CHI 2001, Extended Abstracts*, pp. 269–270. ACM Press (2001)
7. Ayaka, S., Keita, W.: MimiCook: a cooking assistant system with situated guidance. In: *TEI 2014*, pp. 121–124. ACM Press (2014)
8. Yu, S., Shunsuke, M.: Cooking support with information projection onto ingredient. In: *APCHI 2012*, pp. 193–198. ACM Press (2012)
9. Sebastian, B., Markus, F.: Using head-mounted displays and in-situ projection for assistive systems – a comparison. In: *PETRA 2016 (2016)*. Article No.44

10. Markus, F., Tomas, K.: Interactive worker assistance: comparing the effects of in-situ projection, head-mounted displays, tablet, and paper instructions. In: *UbiComp 2016*, pp. 934–939. ACM Press (2016)
11. Marina, C., Stephan, L.: Workspace awareness in collaborative AR using HMDs: a user study comparing audio and visual notifications. In: *AH 2016*. ACM Press (2016). Article No.3
12. Youngsun, K., Seokjun, H.: Augmented reality based remote coaching system. In: *VRST 2016*, pp. 311–312. ACM Press (2016)
13. Nielsen, J., Molich, R.: Heuristic evaluation of user interface. In: *Proceedings of ACM CHI 1990 Conference on Human Factors in Computing System*, pp. 249–256. ACM Press, New York (1990)
14. Nielsen, J.: Usability inspection method. In: *Conference Companion on Human Factors in Computing System*, vol. 25, no. 1, pp. 377–378 (1995)
15. De’Longhi Homepage. <https://www.delonghi.com/zh-cn/>. Accessed 24 Feb 2014
16. Jianming, D., Limin, F.: *Human Computer Interaction*, 5th edn. Tsinghua University Press, Beijing (2016)
17. Tom, T., Bill, A.: *Measuring the User Experience*, 2nd edn. Elsevier, Singapore (2016)
18. National Aeronautics and Space Administration. NASA-TLX Work Load Index. <https://humansystems.arc.nasa.gov/groups/tlx/>. Accessed 27 Apr 2018