The Advantages and Limitations of a Wearable Active Camera/Laser in Remote Collaboration

Takeshi Kurata[1,3] Nobuchika Sakata[3,4] Masakatsu Kourogi[3] Hideki Kuzuoka[4] Mark Billinghurst[1,2]
[1]Human Interface Technology Lab, University of Washington, Seattle
[2]Human Interface Technology Lab NZ, University of Canterbury, Christchurch, NZ
[3]AIST, Japan [4]University of Tsukuba, Japan
kurata@ieee.org, {sakata, kuzuoka}@esys.tsukuba.ac.jp
m.kourogi@aist.go.jp, grof@hitl.washington.edu

ABSTRACT
The Wearable Active Camera/Laser (WACL) allows the remote experts not only to set their viewpoints into the wearers' working place independent of their motion but also to point to real objects directly with the laser spot by controlling it through wireless network. In this paper, we examined how communication patterns differ between the WACL interface and a typical headset interface. The results have implications for improving the WACL interface so as to redress the communication asymmetries by enhancing the visual assist.

Categories and Subject Descriptors
H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*computer-supported cooperative work*

General Terms
Design, Human Factors

Keywords
wearable active camera/laser, head mounted display/camera, gesture, augmented reality

1. INTRODUCTION
Unlike most video-conferencing systems, the focus with wearable collaborative systems is on the real world task space. They are suitable for situations where the user wants to move around the task space rather than stay fixed in one place. We are interested in collaboration between a mobile fieldworker and a remote expert such as a network engineer who has to move around while getting directions from a remote supervisor. We have recently developed a WACL [5] that involves wearing a steerable camera/laser head. The WACL is supposed to be attached around a shoulder so as to realize a hands-, eye-, and head-free I/O device for remote collaboration.

In a previous user study [3], we compared remote collaboration with the WACL interface to that with a typical headset interface comprising a head-mounted display (HMD) and a head-mounted camera in terms of task performance, ease of use, and user preference. By examining task completion time and questionnaire results obtained from a series of Lego block selection and assembly tasks, we found that the WACL gave better impressions to the wearer in terms of comfortability when wearing, eye-friendliness, and fatigue, in spite of no significant difference in the total completion time between the headset and the WACL.

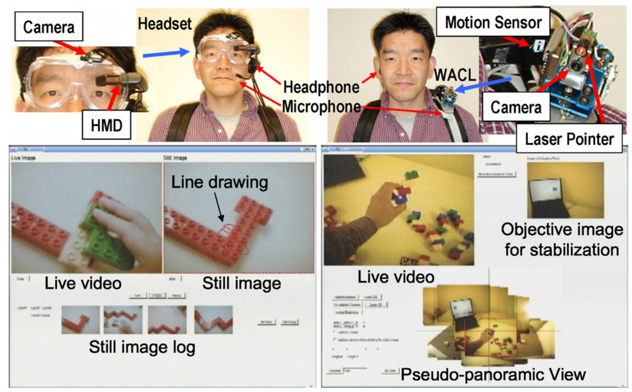

Figure 1: Headset interface and WACL interface.

In this paper, we clarify the advantages and limitations of the WACL interface by examining how communication patterns differ between the WACL and the headset from transcripts of video log data collected in the previous study. As described in [2], task phases during collaboration in real world tasks primarily consist of object/location identification, procedural instruction, and comprehension monitoring. Visual assist by pointing such as cursor pointing on images and laser pointing on real objects is beneficial for the identification phase, but not for the procedural instruction phase. Meanwhile, if there is no object in images taken from a wearable camera, we cannot even use pointing. Thus we hypothesized that remote experts would talk more to fieldworkers wearing the WACL during procedural instruction phase and would talk more to ones wearing the headset when view changes are required.

2. METHOD
2.1 Equipment and Task
Figure 1 shows the headset worn by fieldworkers (upper left), the GUI for experts to interact with the headset wearers (lower left), the WACL (upper right), and the GUI for experts to interact with the WACL wearers (lower right). The headset interface provides users with visual assist not only by pointing, but by line drawing on live/still images. Func-

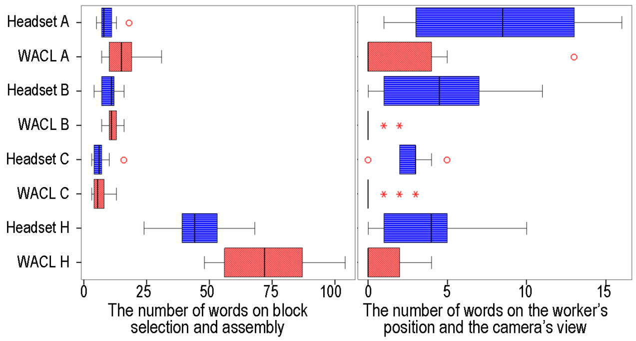

Figure 2: The number of experts' words by section and media condition. This box-and-whiskers plot shows the median, quartiles, and outliers.

tions that the headset interface has are similar to DOVE [4]. In the context of wearable interface, line drawing on still images is simple but effective way to prevent a registration problem between images and the line drawing.

An experimental workplace contained four sections (**A**, **B**, **C**, and **HOME**). Dozens of *block clusters* assembled with several Lego blocks were distributed in sections **A**, **B**, and **C**. In each trial, under guidance from the remote experts, the workers first had to pick up one or two *block clusters* in each of those three sections, and then complete the trial with block assembly at **HOME** section.

2.2 Subjects and Procedure

Sixteen subjects (7 female, 9 male) served as fieldworkers, and two male experts were paired with eight subjects each. Pairs did a trial for training and an actual trial with each of two media conditions. Transcripts of the actual trials were made of video log data manually, and we counted up the number of *morae* (*mora*: phonetic unit in Japanese), words on block selection/assembly, and words on the worker's position and the camera's view by participant role and section.

3. RESULTS

The number of *morae* that experts uttered accounted for about 90% of all in both media conditions. *Morae* of experts were significantly fewer in the headset condition than in the WACL condition (Wilcoxon signed rank test, $p = 0.03$), but there was no difference between two conditions as to workers. Accordingly, we focus here on the expert side. Figure 2 shows the number of words on block selection/assembly and that on position/view by section and media condition. Experts used significantly fewer words on block selection/assembly in sections **A** and **HOME** with the headset condition than with the WACL condition ($p = 0.017$ and 0.002), but there was no difference between two conditions in sections **B** and **C**. In contrast, they used significantly fewer words on position/view in every section with the WACL condition than with the headset condition ($p = 0.01$, 0.007, 0.008, and 0.016 for **A**, **B**, **C**, and **HOME**, respectively).

4. DISCUSSION AND FUTURE WORKS

Although conversational analysis from two experts might not be statistically meaningful, these results agree with our hypothesis. Visual assist by pointing was beneficial for the identification phase in sections **A**, **B**, and **C**, but not that much for the procedural instruction phase in **HOME** section. Since experts could use more advanced visual assist by line drawing in the headset condition, they did not need to explain the details about block assembly in **HOME** section compared with in the WACL condition. Both the experts commented that explaining block assembly imposed more burdens on them in the WACL condition. In section **A**, the resolution of *block clusters* on images had to be larger to identify them than in sections **B** and **C** (see [3]). Since the workers with the WACL had no means of confirming that unlike with the HMD, the experts sometimes sent misdirection to the workers and that increased the number of words on block selection. Meanwhile, the experts had difficulty in observing everything at a glance in every section. In the headset condition, they often needed cooperation from the workers so that they were able to look at the workingplace as they wished, and that increased the number of words on position/view.

It was found that the WACL interface induced several communication asymmetries [1], which imposed more burdens on the experts as described the above and gave better impressions to the workers as reported in [3]. One practical means of redressing the asymmetries is to equip the WACL user with an additional display device for presenting detailed visual assists. A Shoulder-Worn Display (SWD) [5] may be suitable for this purpose since the SWD is supposed not to spoil the advantages of the WACL which is hands-, eye-, and head-free interface. We are currently assessing how well a WACL/SWD condition works compared with the WACL-only case. Another possibility is to use a MEMS mirror for laser scanning to project detailed visual assists directly on the real workingplace.

ACKNOWLEDGEMENTS: This work is supported in part by Special Coordination Funds for Promoting Science and Technology from MEXT and by JSPS Postdoctoral Fellowships for Research Abroad (H15).

5. REFERENCES

[1] M. Billinghurst, S. Bee, J. Bowskill, and H. Kato. Asymmetries in collaborative wearable interfaces. In *Proc. ISWC99*, pages 133–140, 1999.

[2] S. R. Fusell, L. D. Setlock, E. M. Parker, and J. Yang. Assessing the value of a cursor pointing device for remote collaboration on physical tasks. In *Proc. CHI 2003*, pages 788–789, 2003.

[3] T. Kurata, N. Sakata, M. Kourogi, H. Kuzuoka, and M. Billinghurst. Remote collaboration using a shoulder-worn active camera/laser. In *(submitted to ISWC 2004)*, 2004.

[4] J. Ou, X. Chen, S. R. Fusell, and J. Yang. DOVE: Drawing over video environment. In *Proc. MM'03*, pages 100–101, 2003.

[5] N. Sakata, T. Kurata, T. Kato, M. Kourogi, and H. Kuzuoka. WACL: Supporting telecommunications using wearable active camera with laser pointer. In *ISWC 2003*, pages 53–56, 2003.