

# Evaluation of View Angle for a First-person Video to Support an Object-finding Task

**Tatsuyuki Kawamura (kawamura@is.naist.jp)**

**Yasuyuki Kono (kono@is.naist.jp)**

**Masatsugu Kidode (kidode@is.naist.jp)**

Graduate School of Information Science

Nara Institute of Science and Technology

8916-5 Takayama, Ikoma, Nara 630-1092 JAPAN

**Takahiro Ueoka (ueoka@tome.jp)**

Tome R&D Inc.

134 Chudoji Minamimachi, Shimogyo-ku, Kyoto 600-8813 JAPAN

## Abstract

This paper describes an investigation of an effective view angle for a first-person video to support an object-finding task. We conducted experiments where subjects evaluated videos with view angles of 45 through 135 degrees. We found that the useful view angle range for recognizing “the action” and “the location” is from 115 to 125 degrees.

## Introduction

This paper describes an investigation into the influence of view angle on a first-person video to human cognition of “action” and “location” in a video for supporting an object-finding task. Here, a first-person video refers to a video recorded with a wearable camera set on a user’s head.

In this paper, we investigate the appropriate range for a camera’s view angle. This investigation is important for designing a wearable camera for supporting an object-finding task (Ueoka, Kawamura, Kono & Kidode, 2003). Sometimes, the user cannot find an object even though he or she refers to the first-person video including the object (Ueoka, Kawamura, Kono & Kidode, 2004). A segment of a video has a time limitation. This limitation causes the user to not recognize the correct context of the video because an ongoing-action in the video is cut off. The view angle also limits the person’s spatial cognition. It is difficult for the user to recognize the location in the video when the system gives only a small amount of information about the location. A very wide view angle in a camera enables the user to understand both the action and the location. However, a narrower view angle is better than a wide view angle for a miniaturizing and mounting the camera device.

## View Angle Evaluation and Results

Fifteen subjects participated in these experiments. All of the subjects were male students of a graduate school who studied information science in a laboratory.

All experimental videos, which were prepared to be shown to each subject, displayed scenes from an experimenter’s viewpoint recorded by a head-mounted CCD camera worn by the experimenter. The view angle conditions of the videos were virtually simulated with

trimmed source videos. To establish a wide range of view angle conditions, three source videos were recorded with three types of lenses alternately attached to the camera. The source videos were trimmed in a total of ten view angle stages to simulate view angle conditions.

All videos were displayed on a 16-inches LCD at the same magnification. Each source video was recorded when the experimenter selected a positive or negative action by placing a target object, in this case a book, in a place. Ten places in the laboratory where students spent time were selected as target places for the experiment. A positive action represents an experimenter actually placing the target object in a location. In contrast, a negative action represents an experimenter not placing the target object in a location, but first holding on to the object and then bringing it to another location, not designated as a target place. Positive actions were held in five places, half of all the places, and negative actions were held in the other half. The experimenter conducted these behaviors naturally while recording videos.

Each experimental video was displayed on a browser. The interface enabled each subject both to replay an experimental video and to answer a questionnaire. The questionnaire was denoted as follows:

**Q1** Do you think that the target object was really placed there? Vote the following options:

- (a) It was really placed there.
- (b) It was not placed there.
- (c) I have no idea.

**Q2** Do you think where the target object was placed? Mark there on the map if you can think any proposed place, or you should mark on “no idea.”

To answer the first question, each subject had to decide whether the experimenter really placed or did not place the target object on any place displayed in the experimental video. If it was difficult for the subject, he could answer that he had “no idea”. The second question mentions cognitive estimations of each subject from viewing a place displayed in a viewpoint video. The map displays arrangements of

fixed furniture in the laboratory environment, such as walls, columns, doors, bookshelves, desks and copiers. The subject was allowed to answer a number of proposed places up to 30 in all. He was also allowed to answer “no idea” if the subject could not focused on the proposed places. After the subject completed the voting forms, he went on to the next questionnaire. Questionnaires in a trial were ordered by the view angle level of the experimental videos in an ascending sequence. Through all of the experiments for a subject, trials were ordered by the viewing angles of lenses into an ascending sequence so that all experimental videos were ordered into an ascending sequence of the view angle.

From the questionnaires written by the subjects, the results of action recognition and location recognition were counted on a view angle basis. The answers to videos in which sources were taken in locations 3 and 8 are not included in the counted results because of some significant gaps of context between sources with different lenses.

Figure 1 illustrates the subject number ratio of action recognition separated into TRUE-POSITIVE (Figure 1 (a)), and FALSE-NEGATIVE (Figure 1 (b)). TRUE/FALSE means that a subject recognized the action of placing a target object in each video correctly/incorrectly. POSITIVE/NEGATIVE refers to the contents of the subject’s answer of whether the action was actually occurred/did not occur. We found that the ratios of TRUE-POSITIVE increase and FALSE-NEGATIVE decrease with increases in the view angle condition.

Figure 2 (a) illustrates that the true results of location recognition were saturated from the 115-degree view angle condition. The false results illustrated in Figure 2 (b) almost decreased to 20% when the view angle was set to 105.

### Discussion

Using the experimental results, we defined the required and sufficient view angle conditions to support a user’s object finding task. When the view angle was 115, the true graph of location recognition (Figure 2 (a)) was saturated with the false graph, and the candidate location graph of location recognition (Figure 2 (b)) was asymptotically stable at each value. From these results, the required view angle condition should be set to 115. On the other hand, the sufficient view angle condition should be set to 125, because the true positive graph of action recognition was saturated at this setting.

### Concluding Remarks

In this paper, we have investigated the effective view angle of the lens for a first-person video. Through experiments with subjects, we found that a lens with a range of view angle from 115 to 125 (degrees) is better suited for recording a first-person video.

### Acknowledgement

This research was partially supported by Core Research for Evolutional Science and Technology (CREST) Program “Advanced Media Technology for Everyday Living” of

Japan Science and Technology Corporation (JST). Also, this research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists (B), 17700113, 2005.

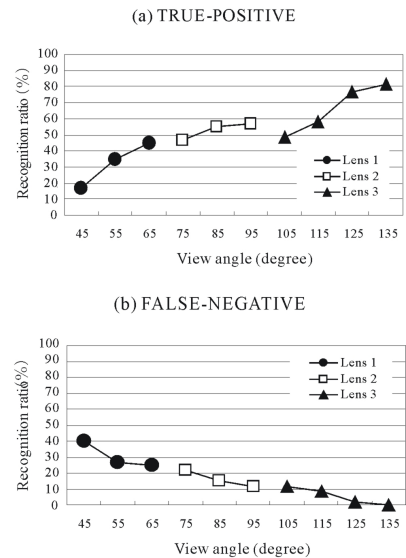


Figure 1: Results of action recognition

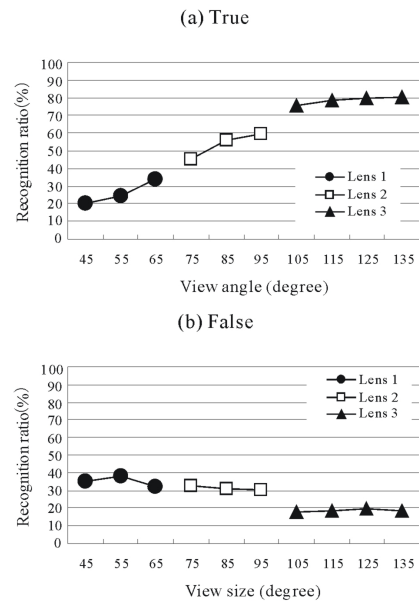


Figure 2: Results of location recognition.

### References

Ueoka, T., Kawamura, T., Kono, Y., & Kidode, M. (2003). I’m Here!: a Wearable Object Remembrance Support System. *Proceedings of the Fifth International Symposium on Human Computer Interaction with Mobile Devices and Services* (pp.422-427).

Ueoka, T., Kawamura, T., Kono, Y., & Kidode, M. (2004). Evaluation of a Vision-based Object Remembrance Support System. *Proceedings of the IEEE International Conference on Multimedia and Expo*.