PERCEPTION OF MULTI-RESOLUTION IN STEREOSCOPIC IMAGES

Masahiko Ogawa[1], Kazunori Shidoji[2], and Yuji Matsuki[2]
[1]*Graduate School of Information Science and Electrical Engineering,*
[2]*Faculty of Information Science and Electrical Engineering,*
Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
<{ogawam, shidoji, matsuki}@brain.is.kyushu-ac.jp>

Abstract

Images that are high resolution and wide angle cannot be acquired with a camera. Due to characteristics of the human visual system, however, the entire wide-angle image does not necessarily need to be high resolution. Thus, we examined the minimum resolution required in the field of view. The results showed that observers perceived the image similarly when the entire image was high resolution and when the resolution of the image at positions that were more than 20 and 40 degrees from the center of the visual field decreased to 25% and approximately 11% of the resolution of the gaze point, respectively. Additionally, we investigated whether the participants were able to distinguish between the original completely high-resolution image and processed images, which included triple-resolution, dual-resolution, and low-resolution images. The results demonstrate that the participants could not discriminate the triple-resolution images from the original images. Finally, we developed a stereoscopic camera system based on these results.

Recent developments in virtual reality technology have resulted in systems that can result in highly realistic sensations. A camera and monitor system that projects the real world as it is, however, has yet to be developed. This is because of a technical limitation; high-resolution and wide-angle images are not compatible.

There are two ways to solve this problem. One is to develop high-resolution digital television systems, such as the ultrahigh-definition color video camera (Shimamoto *et al.*, 2005). The other is to combine existing systems. Here we have attempted to employ the latter approach.

Dual resolution stereoscopic video systems have been proposed as teleoperation systems (Matsunaga *et al.*, 1998; Shidoji *et al.*, 2003; Ienaga *et al.*, 2005). In these stereoscopic systems, high-resolution images that were taken with telephotographic cameras were inserted into low-resolution images taken with pantoscopic cameras and displayed at their full sizes. It was reported that the users of these systems were able to operate the machines more efficiently than when they used conventional monoimage systems. The observers were able to see low-resolution images in the peripheral area of the visual field, because these systems were not designed to improve reality of images.

Naemura *et al.* (2000) proposed a quasi-tri resolution stereoscopic system. This system included dual-resolution images that had different high-resolution areas for right and left eyes. The authors reported that when they enlarged the central high-resolution area for the left eye, the observers had a difficult time recognizing low-resolution objects in the peripheral of the visual field.

These results suggested that observers can perceive superimposed multi-resolution images as a single completely high-resolution image if the images are properly overlapped based on characteristics of the human visual system; *i.e.*, individuals have high visual acuity in the central field of view, whereas the visual acuity in the peripheral field is low. We examined the compositions of multi-resolution stereoscopic images that observers could not distinguish from completely high-resolution images (Ogawa *et al.*, 2007). The results showed that observers perceived the images similarly when the completely high-resolution images were presented with images that contained areas in which the resolution was 25% of the original resolution at positions more than 20 deg outside of the central position. Moreover, in areas shifted by more than 40 deg from the gaze point, we surmised that a resolution that was approximately 11% of the resolution at the gaze point was sufficient. We also suggested that observers were barely able to perceive decreased resolution in the peripheral visual field of triple-resolution images in which the resolution varied in three separate areas.

In this study, we experimentally investigated whether the previously proposed triple-resolution images and the completely high-resolution images are perceived similarly. We then describe a newly developed, multi-resolution stereoscopic image system.

Method

Participants

Seven participants participated in this study, all of whom were male university students between the ages of 22 and 25 years old. They all had normal or corrected-to-normal acuity.

Apparatus and stimuli

A Wheatstone stereoscopic viewer using monitors and mirrors was used for stimulus presentation (Figure 1). Stimuli were computer graphic images generated using a personal computer (CPU: Intel Xeon 3.06 GHz; Graphic board: NVIDIA Quadro FX3000) and displayed on two LCD monitors (iiyama: AQ5311D-BK). These monitors were set up facing each other. Between the monitors, mirrors were set at a 45-deg angle. At a viewing distance of 450 mm, these monitors were subtended 54-by-41 arc deg.

Eye movements were monitored using a video camera (Sony: DCR-HC41) sampling at a rate of 30 Hz. The video stream was recorded using another personal computer (CPU: Pentium 4.3 GHz). Gaze points were measured using image processing of the pupil and a HALCON 7.1 software library (MVTec).

There were four different stimuli, each of which included a pair of images with binocular disparity. The original high-resolution image (resolution: 2048 × 1536 pixels) was created using OpenGL. A red cross was drawn at the center of the image to provide the gaze point. In the image, a number of small 2-cm diameter blue spheres were shown at equal intervals in a 13 × 17 array. The spheres were alternately presented in anterior and posterior positions to provide depth perception to the observers. The even-numbered spheres on the even-numbered rows and the odd-numbered spheres on the odd-numbered rows were located 47 cm from the observers. The remaining spheres were approximately 45 cm from the observers. Each sphere occupied approximately 2.8 deg of visual angle for the participants.

The other three stimuli, which were triple-resolution, dual-resolution, and low-resolution images, were created by processing the original high-resolution image using Photoshop Elements (Adobe). The nearest neighbor method was used as the algorithm to change the resolution. The resolution of the center area, which subtended 16 × 12 deg in the triple-resolution image, was the same as that in the high-resolution image. The pixel density of the area subtended between 16 × 12 deg and 32 × 24 deg was decreased to 25% of that of the center area. The pixel density of the remainder of the visual field (from 32 × 24 deg to 54

× 41 deg) was decreased to approximately 11% of that of the center area. In the dual-resolution image, the pixel density of the area outside of the center area subtended within 16 × 12 deg was decreased to approximately 11% of that of the center area. In the low-resolution image, the pixel density was decreased to 11% of that of the high-resolution image throughout the image (640 × 480 pixel).

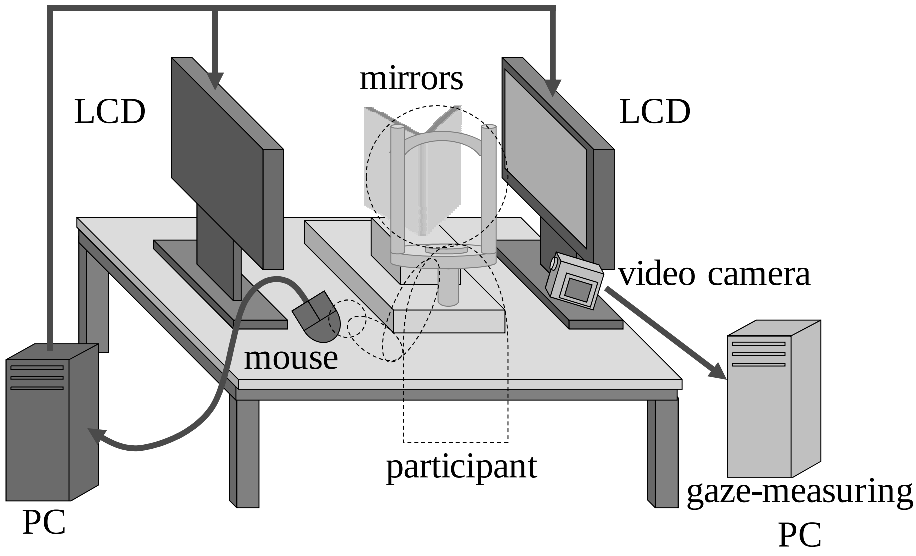

Fig. 1. The Wheatstone stereoscopic viewer that was used to present the images with binocular disparity to each participant's eyes. The heads of the participants were fixed using a head restraint.

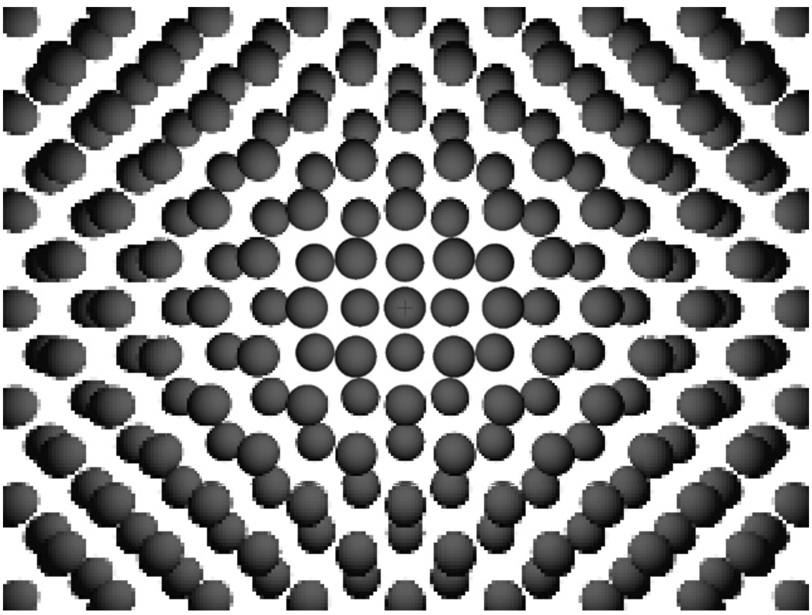

Fig. 2. A reproduction of the triple-resolution image in which the resolution was decreased in the two outer areas. The central area subtended 16 × 12 deg had the same resolution as the original high-resolution image (2048 × 1536 pixels). The middle resolution area subtended between 16 × 12 deg and 32 × 24 deg had a pixel density that was 25% of that of the central area (1024 × 768 pixels). The outer low-resolution area had a resolution that was the same as the low-resolution image (640 × 480 pixels).

Procedure

The experiment was performed in a dark room. The original high-resolution stereoscopic images were displayed first. Each participant placed their head in the restraint to view the images and the fixated on the central red cross in the image with both eyes. When the participants pressed the mouse button, one of four stimuli was randomly selected and presented to the subject for 1 sec. Then the participant determined whether the two images were the same or not, and pressed a corresponding button on the mouse. Each stimulus was presented 30 times.

Results and Discussion

Using a video image of the pupils of each participant, some small eye movements corresponding to pixel displacements of approximately 10 pixels were observed. When the participants shifted their gaze from the central sphere in the image to an adjacent sphere, their pupils moved approximately 15 pixels on the video image, which corresponded to a 2.5-deg movement of the gaze point. Thus, we excluded trials in which the pupils of the subject were displaced by more than 15 pixels. About 5% of the data were excluded due to this criterion.

Figure 3 shows the mean answer rates for which the participants perceived that the two images were the same. The mean answer rates were 0% for the low-resolution trials, 4% for the dual-resolution trials, 77% for the triple-resolution trials, and 93% for the high-resolution trials. One-way ANOVA with repeated measures was used to analyze the results from the various trials. There was a significant main effect from the type of trial [$F(3, 18) = 81.039$, $p < .001$]. Multiple analyses using the Ryan method demonstrated that the answer rate for the low-resolution image was significantly lower than those for the triple-resolution image [$t(3) = 10.114$, $p < .05$] and the high-resolution image [$t(4) = 12.278$, $p < .05$]. Furthermore, the answer rate for the dual-resolution image was significantly lower than those for the triple-resolution image [$t(2) = 9.545$, $p < .05$] and the high-resolution image [$t(3) = 11.708$, $p < .05$]. There were no significant differences between the results obtained for the other stimuli.

The lack of a significant difference between the perecptions of the triple-resolution image and the high-resolution image indicates that the participants did not perceive the decrease in the peripheral resolution in the triple-resolution image. This may have resulted from the middle area in the triple-resolution image, which smoothed the transition from the high-resolution area to the low-resolution area.

Although we still need to examine a number of variables, such as the spacial frequencies, contrasts, colors, and objects used for the stimuli, the results from this experiment suggest that we can compress image data without reducing information using the triple-resolution configuration. In this experiment, the number of pixels in the triple-resolution image was 921,600 (640 × 480 × 3), whereas there were 3,145,728 pixels (2,048 × 1,536) in the completely high-resolution image. Thus, the amount of information in the former was about 29% of that in the latter.

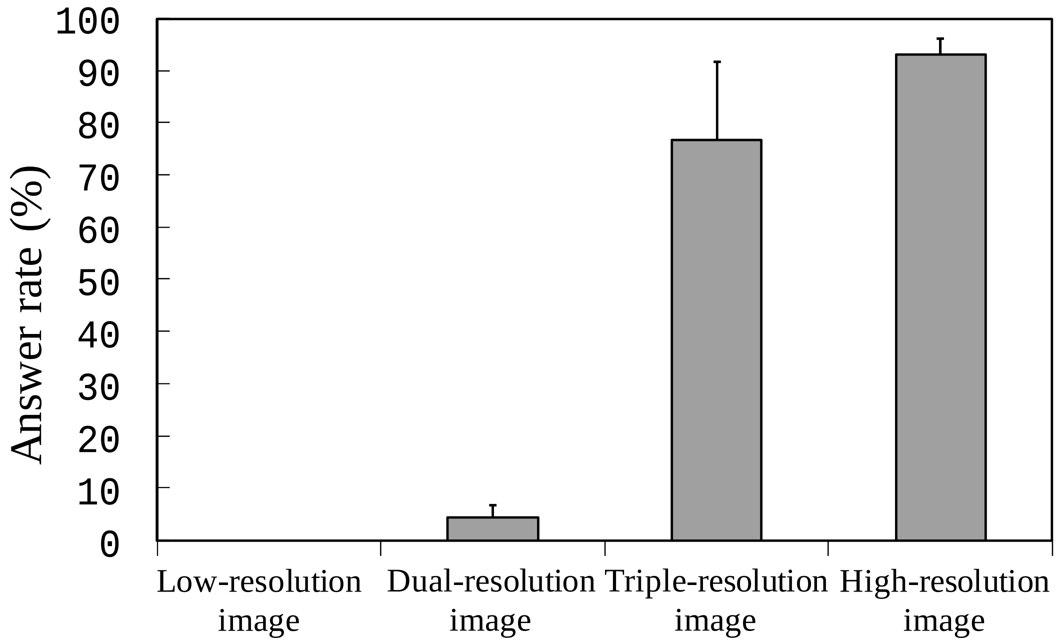

Fig. 3. Mean answer rate after viewing the low-resolution, dual-resolution, triple-resolution, and high-resolution images.

A new triple-resolution stereoscopic system

We then developed a triple-resolution stereoscopic camera system based on our results (Figure 4). Images were obtained for left and right eyes using three IEEE1394 cameras (Point Grey Research: FLEA-HICOL-CS) to create a stereoscopic view. All of the obtained images contained 640 × 480 pixels (VGA). For each eye, the cameras for the high- and middle-resolution areas had varifocal lenses (Omron: 13VM550T), the angles of which were set at 21 and 42 deg. The optical axes of these cameras were matched using a half mirror. The camera for the low-resolution area was attached to a wide lens (Omron: 13FM28IR), the angle of which was 68 deg. These camera angles were based on the results from the experiment and the settings of the Wheatstone stereoscopic viewer.

Triple-resolution video images were created as follows. The three video streams were captured using personal computers for the left eye (CPU: Intel Xeon 3.06 GHz; Graphic board: NVIDIA Quadro FX3000) and the right eye (CPU: Intel CORE2 6600 @ 2.40 GHz; Graphic board: NVIDIA GeForce 6600GT). The display sizes of the images for the middle- and low-resolution areas were enlarged to 1280 × 960 and 2048 × 1536 pixels, respectively. Thus, the pixel densities of the images were decreased to 25% and approximately 11% of the original density, respectively. Finally, these streams were made to overlap using image processing.

Participants viewed the triple-resolution images through the Wheatstone stereoscopic viewer. If they did not move their eyes, the participants reported that they did not perceive a decrease in the peripheral resolution in the triple-resolution image. In further studies to improve the system, the cameras will be moved according to the participant's gaze point using a camera positioner.

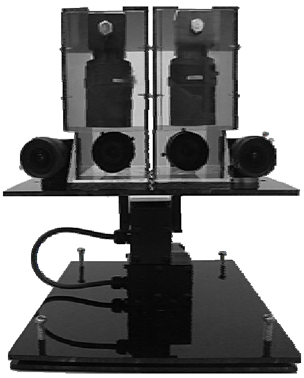

Fig. 4. The stereoscopic camera system used to obtain the triple-resolution images.

Acknowledgements

This study was supported in part by a Grant-in-Aid for Scientific Research (c) from the Ministry of Education, Culture, Sports, Science and Technology (project number 17500071), and by the 21st Century COE Program "Reconstruction of Social Infrastructure Related to Information Science and Electrical Engineering."

References

Ienaga T., Matsunaga K., Shidoji K., Otsuru M., Araki S., & Matsuki Y.(2005). An effect of large overlapped area of stereo pairs at the working point on a spatial multi-resolution stereoscopic video system, *Proceedings of IEEE Virtual Reality 2005*, pp.277-278.

Naemura, T., Sugita, K., Takano, T., & Harashima, H. (2000). Multi-Resolution Stereoscopic Immersive Communication Using a Set of Four Cameras, *Proceedings of Stereoscopic Displays and Applications XI of the International Society for Optical Engineering: Stereoscopic Displays and Virtual Reality Systems VII*, Vol.3957, pp.271-282.

Matsunaga, K., Nose, Y., Minamoto, M., Shidoji, K., Ebuchi K., Itoh, D., Inoue, T., Hayami, T., Matsuki, Y., Arikawa, Y., & Matsubara, K. (1998). A new stereoscopic video camera and monitor system with central high resolution, *Proceedings of 10th International Symposium of the International Society for Optical Engineering, Stereoscopic Displays and Virtual Reality Systems V*, pp.164-170.

Ogawa, M, Shidoji, K., & Matsuki, Y. (2007). Development of multi-resolution stereoscopic image system. *Report of fostering program of young researchers 2006 (the 21st Century COE Program "Reconstruction of Social Infrastructure Related to Information Science and Electrical Engineering")*, pp.122-126.

Shidoji, K., Matsunaga, K., Minamoto, M., Nose, Y., Ebuchi, K., & Matsuki, Y. (2003). Visual interface for remote control. *Journal of Robotics and Mechatronics*, Vol.12, No.1, pp.40-46.

Shimamoto, H., Yamashita, T., Koga, N., Mitani, K., Sugawara, M., Okano, F., Matsuoka, M., Shimura, J., Yamamoto, I., Tsukamoto, T., & Yahagi, S. (2005). An 8k x 4k Ultrahigh-Definition Color Video Camera with 8M Pixel CMOS Imager. *SMPTE Motion Imaging Journal*, pp.260-268.