RECOVERY OF DEPTH FROM VIDEO IMAGE

Toshio Watanabe and Keiji Nishioka
Department of Environmental Information, Keio University,
5322 Endoh, Fujisawa 252-0816 Japan
E-mail: watanabe@sfc.keio.ac.jp knis@sfc.keio.ac.jp

Abstract

How can we reproduce depth accurately from a photograph or a video image of 3D scene? In the present study subjects observed 3D scene in three conditions (video image observation, direct observation and stereoscopic observation). Subjects were asked to judge the distance from the origin to an object in the scene and the angle of the object from the sagittal line. As a result, we found that both the distance and the angle of the object in the video image were non-linearly related to those in the direct observation. And we found that the distance in the stereoscopic observation was almost equal to that in the direct observation, but the angle was non-linearly related to that in the direct observation.

How can we reproduce depth accurately from a photograph or a video image of 3D scene? It is known that there are systematic distortions in depth judgments in photographs of natural scenes (Smith 1958; Kraft, Patterson, and Mitchell 1986). Recently Hecht et al. (1999) studied flattening effects of angle judgment and distance distortions in photographs. They took photographs of building corners in a natural scene. The subject was asked to get an impression for the angle subtended by the two walls forming the corner (16.5,67,72,108,113 and 161.5 degrees of angle) and for the distance to the corner (1.5m,10m and more or less). As a result the angles were overestimated for near and far viewing distances, and the distances were overestimated for near viewing distance and underestimated for far viewing distance. In the present study three experiments were conducted to find the quantitative relation among the depth in the video image observation, the depth in the direct observation and the depth in the stereoscopic observation. In the video image observation the subject observed a three dimensional scene through the video image projected onto the screen. In the direct observation he observed the scene directly and in the stereoscopic observation he observed the scene in stereoscopic vision. By comparing three conditions we will find the accuracy of depth perception through the video image and the stereoscopic vision.

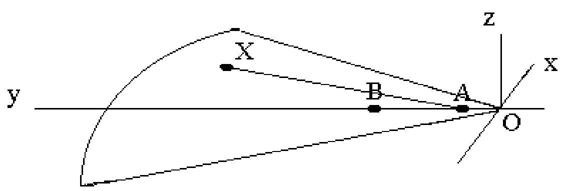

Figure 1. The objects A,B and X placed in the horizontal plane. The subject or the videocamera was located at O.

Method

In the video image observation (condition 1), as shown in Figure 1, three objects A, B, X were placed in the horizontal plane. The objects were 3cm in diameter and 1cm in height. The experimental setting including three objects was recorded by a digital video camera (Victor Digital Video Movie GR-DVX, 27-mm lens) placed at O and was projected onto the screen by the digital projector (Plus U2-870, 27.5-mm lens). The lens position of the video camera was 12cm above the horizontal plane. The distance between the screen and the projector was 280cm. Subject was seated in back of the projector and asked to judge distance AX and angle BAX as distance AB =1. One trial consists of a pair of judgments of the distance and the angle of object X. Subject's eye level was almost the same as the height of the projector. Physical distance AX was one of 0.4m,1.0m,1.6m,2.2m,and 2.8m, and physical angle BAX was one of ±3, ±8, ±13,

±18,and ±23 degrees of angle. Positive angles show the right side of the sagittal line and negative angles show the left side. Total locations of object X were 25 locations, 13 locations for the right side and 12 locations for the left side of the sagittal line. In each trial one location was randomly selected from 25 locations. Each condition consists of 25 trials. Sixteen subjects (8 male and 8 female) participated in condition 1. They observed the video image binocularly. In the direct observation (condition 2), the video camera was removed and the subject observed three objects A, B and X at O to judge distance AX and angle BAX. The subject's eye level was 12cm above the horizontal plane and it was the same as the height of the lens of the video camera. Twelve subjects (6 male and 6 female) participated in condition 2. In the stereoscopic observation (condition 3), two video cameras were placed side by side at O and each video camera was connected to the corresponding digital projector. Two video images were projected to the screen simultaneously and one image was polarized vertically and the other horizontally. The subject observed the screen at the same location of condition 1. The subject wearing polarized glasses judged distance AX and angle BAX. The remaining procedures were the same as in condition 1. Twelve subjects (6 male and 6 female) participated in condition 3. All subjects participated in only one condition. They had normal vision or corrected-to-normal vision, and they were naive with respect to the purpose of the study.

Results and Discussion

Figure 2 shows the relationship between the visual distance obtained in the video image observation (δ_1) and the visual distance obtained in the direct observation (δ_2). Each distance was averaged over subjects and over 5 directions (±3,±8,±13,±18,and±23 degrees of angle) in each of conditions 1 and 2. The symbol o shows the relationship between δ_1 and δ_2 , and the symbol + shows the relationship between δ_1 and the corresponding physical distance. As shown in the figure, the symbol + is almost on the solid line. It means that δ_1 is almost equal to the corresponding physical distance. Further, the symbol o seems to be deviated from the solid line. It means that δ_1 may not be equal to δ_2 . By curve fitting, we get

$$\delta_2 = 0.852\,\delta_1^{1.20} \qquad (1)$$

The symbol x in the figure shows the theoretical value of δ_2 obtained by the function (1).

The coefficient of determination is 0.998. The explained variance by the function (1) is significantly larger than the one by the function $\delta_2 = \delta_1$ ($F(4,4)=8.113, p=0.03$). This suggests that δ_1 is longer than δ_2 in short distance, but the former is shorter than the latter in long distance. And it supports the previous study (Hecht et al. 1999)

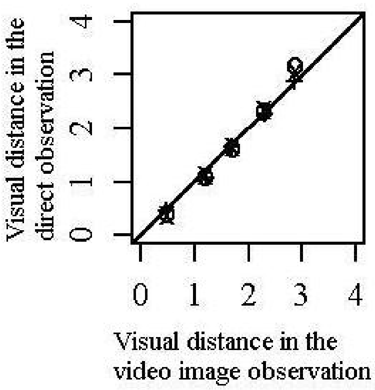

Figure 2. Visual distances in the video image observation and in the direct observation

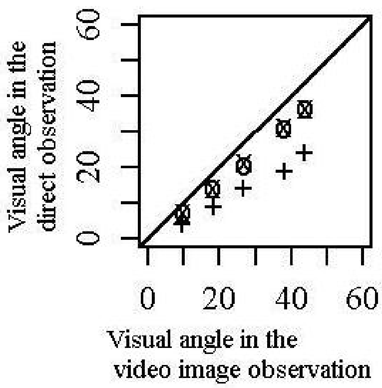

Figure 3. Visual angles in the video image observation and in the direct observation

Figure 3 shows the relationship between the visual angle obtained in the video image observation (φ_1) and the visual angle obtained in the direct observation (φ_2). Each angle was averaged over subjects and 5 distances (0.4m, 1.0m, 1.6m, 2.2m and 2.8m) in each of conditions 1 and 2. The symbol o shows the relationship between φ_1 and φ_2, and the symbol + shows the relationship between φ_1 and the corresponding physical angle. As shown in the figure, φ_1 is larger than both φ_2 and the corresponding physical angle. By curve fitting, we get

$$\varphi_2 = 0.538\,\varphi_1^{1.11} \qquad (2)$$

The symbol **x** in the figure shows the theoretical value of φ_2 obtained by the function (2). The coefficient of determination is 0.998. The explained variance by the function

(2) is significantly larger than that by the function $\varphi_2 = \varphi_1$ ($F(4,4)=18.903, p=0.007$). This means that φ_1 is significantly different from φ_2. These results show that depth perception through video image does not reproduce the depth perception in the direct observation. As Hecht et al. indicate, the flattening effects of angles were found also in the present study.

Figure 4 shows the relationship between the visual distance obtained in the stereoscopic observation (δ_3) and the visual distance obtained in the direct observation (δ_2). As shown in the figure, δ_3 seems to be almost equal to δ_2. By curve fitting, we get

$$\delta_2 = 0.878\,\delta_3^{1.13} \tag{3}$$

The coefficient of determination is 0.996. However, the explained variance by the function (3) is not significantly larger than that by the function $\delta_2 = \delta_3$ ($F(4,4)=2.84, p=0.168$). This supports that the stereoscopic vision reproduces the depth perception in the direct observation in distance judgment.

Figure 5 shows the relationship between the visual angle obtained in the stereoscopic observation (φ_3) and the visual angle obtained in the direct observation (φ_2). As shown in the figure, φ_3 seems to be equal to φ_2. By curve fitting, we get

$$\varphi_2 = 1.21\,\varphi_3^{0.93} \tag{4}$$

The coefficient of determination is 0.999. The explained variance by the function(4) is significantly larger than that by the function $\varphi_2 = \varphi_3$ ($F(4,4)=9.91, p=0.023$). This means that stereoscopic vision does not reproduce the depth perception in the direct observation in angle judgment. But, comparing with the angle judgment in the video image observation, the reproduction of depth perception in stereoscopic observation is better than that in the video image observation.

The depth perception through the video image does not reproduce the depth perception in the direct observation. Even the stereoscopic vision, the reproduction of the depth perception was not accurate enough. In order to more accurate reproduction of depth perception from the video image or from the stereoscopic vision we need to transform the original video image. By appropriate transformation, we will be able to reproduce the depth perception in the direct observation. The mathematical functions obtained in the present study indicate that the functional relation between the video image and the direct observation is not linear. This suggests that the transformation should be non-linear.

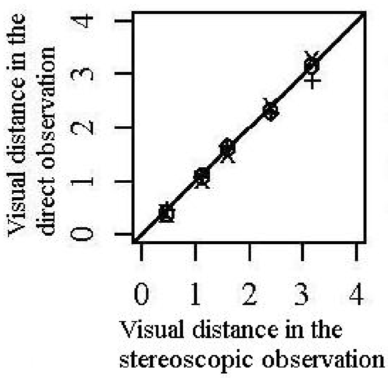

Figure 4. Visual distances in the stereoscopic observation and in the direct observation

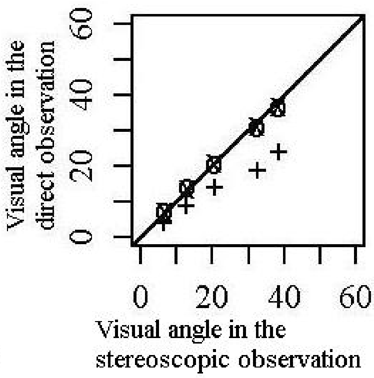

Figure 5. Visual angles in the stereoscopic observation and in the direct observation

References

Hecht, H., van Doorn, A.. and Koenderink, J. J. (1999) Compression of visual space in natural scenes and in their photographic counterparts. *Perception & Psychophysics*, **61**, 1269-1286.

Kraft, R. N., Patterson, J. F. and Mitchell, N. B. (1986) Distance perception in photographic displays of natural settings. *Perceptual & Motor Skills*, **62**, 179-186.

Smith, O. W. (1958) Judgments of size and distance in photographs. *American Journal of Psychology*, **71**, 529-538.