31 augustus 1995

Getting round with fewer views: The effect of a small number of selectable viewpoints on solving transparent knots

W. Pasman, G. J. F. Smets, P. J. Stappers

Delft University of Technology, Faculty of Industrial Design Engineering
Jaffalaan 9, 2628 BX Delft, the Netherlands
W.Pasman@io.tudelft.nl; (+31)15 781791

ABSTRACT

For many practical applications that use camera images and active viewpoint selection to create telepresence it is important to restrict the number of selectable viewpoints. Previous experiments showed that the Delft Virtual Window System can be adapted to work with a small number of selectable viewpoints. For simple scenes and a connectivity task, only three viewpoints proved sufficient, although more viewpoints can compensate for low resolution and contrast. The presented experiment shows that for complex scenes consisting of knots more viewpoints result in better judgements (up to 90% correct) in less time.

KEYWORDS: Visual perception, Interactive displays, 3D displays, fishtank VR, telepresence, delft virtual window system, viewpoint selection, x ray luggage inspection.

INTRODUCTION

Head-slaved displays, such as fishtank VR [10] and the Delft Virtual Window System (DVWS, described below) adapt the image on the screen to the eye position of the observer. This gives the observer a 3D impression. These types of display work well for spatial teleinspection: they are fast to learn and work with, and users make few mistakes. The advantages of such systems in comparison with immersive VR are the lower cost and better integration with the real world [1,2,7,11,14].

All parts of these systems can benefit from or are constrained to a small number of viewpoints:

1. The scene. In some situations only a limited number of viewpoints can be inspected. For example in luggage inspection, each image exposes the luggage to an extra x-ray dose, and the total dose should be as low as possible.
2. The image buffer. In order to be able to respond quickly to changes of viewpoint it may be necessary to store the selectable images in a limited memory. For example 3-D interfaces with pre-calculated viewpoints can use an image buffer.
3. The cameras. The system may use a number of cameras located on the selectable viewpoints. But cameras are expensive, need space and wirings, hinder other movements in the scene and catch away the light. For example the system of Katayama et al [16] works with 7 cameras.
4. The camera positioning apparatus. In stead of using multiple cameras, the system can use one camera and a mechanical construction to move it to the required position. This may allow for a less flexible mechanical construction that can realise only a few viewpoints, resulting in a faster response and a cheaper and lighter construction.
5. The user interface. By lowering the number of selectable viewpoints, the system can use a less sophistocated head tracker that gives only a rough approximation of the actual head position.

Many tasks can be solved with a limited number of viewpoints. For example tree tracing can be done successfully with 11 viewpoints [6], and connections between objects can be judged with only 3 viewpoints [9].

Previous work

Earlier experiments showed that observers using the DVWS are better in performing a depth alignment task than passive observers [1,2]. [5] used a task in which subjects had to decide whether two transparent objects are connected by a wire. An adapted DVWS with a small number of selectable viewpoints proved sufficient. For this task and the DVWS, an interaction exists between image resolution, number of grey levels and the number of selectable viewpoints. for the same task [9] showed that a front, left and right view are sufficient, provided that the resolution and number of grey levels are high enough. It was useful to be able to look $\pm 90^{\circ}$ around the objects, for example to look around an obstacle. Another experiment [7] compared active viewpoint selection to a static viewpoint. It used a knot solving task, similar to the experiment below. The number of correctly solved knots increased from 30% with a static viewpoint to 60% with active viewpoint selection in x-, y- and z-direction. The response time was about 20 s. regardless of the condition. A manipulation task did not have advantage of active viewpoint selection. An explanation is that the observers selected the viewpoint with their head movements, and they did not want to move during manipulation tasks as body movement would interfere with the manipulation.

Some of our experiments [5,9], including the present one, use transparent objects because we are working on an implementation using the DVWS for luggage inspection. This inspection is done with see-through x-ray images.

THE DELFT VIRTUAL WINDOW SYSTEM

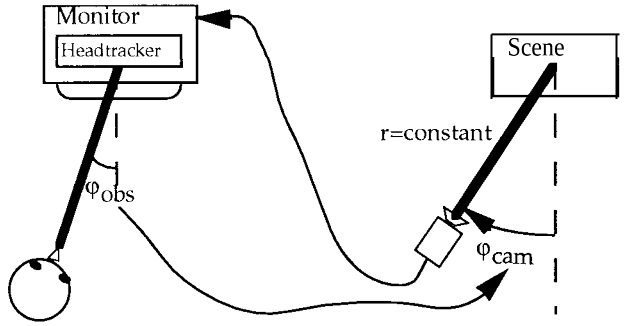

Figure 1: the DVWS in top view. The camera is slaved to the eye positions of the observer.

The DVWS is described in detail in [1,2]. It consists of a monitor, a head tracker, a camera and a scene. Figure 1 shows the system in top view. The monitor displays the image from the camera. The camera rotates in a circle in the horizontal plane around the scene, and is kept aimed at the centre of rotation. The camera rotation is slaved to the rotation of the observer around the middle of the screen.

The Adapted DVWS

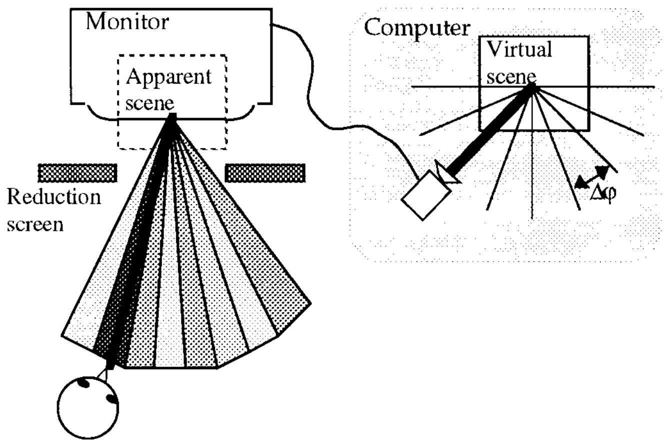

Figure 2: Top view of the adapted DVWS. The camera has 180° freedom of movement and an angle $\Delta\varphi$ between two selectable viewpoints. Each segment of eye position (left) corresponds to a single camera position (right). Observer movements are scaled 4 times.

The adapted system consists of a restriction of the camera movements to the horizontal arc, and of the number of selectable viewpoints. In order to limit the number of selectable viewpoints the camera positions are restricted to discrete steps, as shown in Figure 2. The angle between two selectable camera positions $\Delta\varphi$ was manipulated in the experiment. Because the total freedom of movement of the camera was 180°, the number of selectable viewpoints is $180°/\Delta\varphi + 1$.
Observer movements were scaled four times (i.e. $\varphi_{cam} = 4 \cdot \varphi_{obs}$) to enable the observer inspect a side view. The scene and camera are virtual (i.e. simulated with a graphics package). A reduction screen was placed in front of the monitor, to enhance the perceived depth by reducing the cues for image flatness.

Experimental Hypotheses

In previous experiments [5,9] (see under Previous work) it was expected that large angles $\Delta\varphi$ between selectable viewpoints would disturb the 3D impression and worsen the performance of the operators. However, the angle $\Delta\varphi$ showed no effect when the static image quality was high enough. This suggests that no 3D impression was needed for that task. It is expected however that observers need a 3D impression to judge more complex scenes. The present experiment will use a more complex scene to test whether a large $\Delta\varphi$ worsens observer performance.
In the adapted DVWS there are conflicts between 2D and 3D cues. For example, the focal distance is incorrect for the parts of the scene that are not in the display plane. Furthermore observers using both eyes get the same image for both eyes, causing a potential conflict between the relative shifts of the objects while moving and stereoscopic cues. In order to test whether this last conflict has important effects, observers using one and two eyes will be compared.
Concluding, both large angles between selectable viewpoints and the conflicts with 2D cues are expected to distort the 3D impression of the observer. Therefore, they are expected to increase the response time and decrease the number of correct judgements of the observer.

METHOD

Stimuli

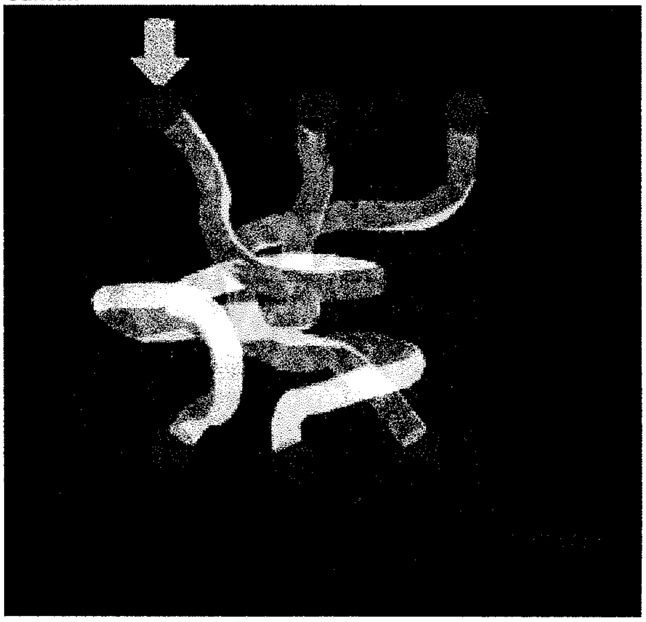

Figure 3: example of a knot. The arrow indicates the start of a wire. Observers had to find the bottom end of that wire.

Each stimulus consisted of three intertwined transparent wires (Figure 3). Each wire started at the top of the screen, and ran through a knot to one of the endpoints at the bottom of the screen. The endpoints are at the left, middle or right when viewed in front view. A red arrow indicated the starting point of one of the three wires. The subjects had to indicate the corresponding bottom end. There were 10 knots for the training and 40 knots for the experiment.

Apparatus

A Dynasight from Origin Instruments was used to track the position of a small reflector that was mounted on a spectacle frame. For the rendering of the image a Silicon Graphics RE Crimson was used. This combination was able to update the image on the screen with 37Hz. The screen had a resolution of 1280x1024 and a size of 33.5 x 28.0 cm. The reduction screen was placed 12 mm in front of the monitor, reducing the image to 22.6 x 17.6 cm. To set the camera in a reasonable perspective projection, the average distance between the head of the observer and the screen was estimated to be 45 cm.

Variables, subjects, design

The independent variables are the angle between two camera viewpoints ($\Delta\varphi$) and the number of eyes used (E). The dependent variables are correctness of choice (C) and the response time (T). Table 1 shows the variables and their levels. Because the total camera range was always 180°, the number of viewpoints (N) was $\frac{180^\circ}{\Delta\varphi} + 1$.

Variable	Levels
$\Delta\varphi$ (N)	5.625” (33), 11.25” (17), 22.5” (9), 45” (5), 90˚ (3)
E	one eye, two eyes
C	correct. incorrect
T	>0 s.

Table 1: variables and their levels

The subjects were 20 students from the Faculty of Industrial Design Engineering (6 women, 14 men). They were naive and paid volunteers. In total, each subject participated about 25 minutes, and got DFl 7.50 for it.
Each subject judged the same 40 knots, but in randomised order and conditions. The conditions were randomised over the subjects in such a way that each knot had been judged once in each condition after five subjects in the same viewing condition. The starting point for each knot (Left, Right or Middle) was randomly chosen. At random, 10 subjects were drawn to view with one eye, the other 10 viewed with two eyes.

Procedure

The subjects in the one-eye condition were asked or tested for their dominant eye. They wore a spectacle frame with a Dynasight reflector and a patch covering the other eye. The subjects in the binocular condition wore the spectacle frame with only a small reflector. Of observers using both eyes the middle of the two eyes was tracked, and of observers using one eye the eye that was used.
The subjects had to follow the indicated wire to the other end. Depending on where it ended, they had to press the left, middle, or right button on a button box below the screen. During the trials they held their hands near the button box. They were asked to test at the start of each trial whether the coupling is of use for them by moving their head. This was done because a pilot study showed that a stimulus with large $\Delta\varphi$ tends to demotivate subjects from moving in subsequent stimuli.
The subjects had 10 seconds to inspect each knot, after which the screen turned dark. A beep was sounded after 8 seconds to warn for the time limit. They were instructed to choose as well as possible, and not to hurry. They could think as long as they wanted, even after the screen went dark. Between the choice and the start of the next trial there was one second delay.
During the training subjects had to judge 10 knots. Immediately after pressing a button, they were informed if they made the right choice, and how long it took to make the choice. During the experiment subjects had to judge 40 knots. Now they did not get feedback, but they were told that they would be told their score after the experiment.

Results

Subjects are enthusiastic about this system, especially if they have a ‘smooth’ image movement (small $\Delta\varphi$). There was no learning effect during the experiment, so the short training was sufficient.
We define the correct-ratio as the fraction of the choices of the subjects that was correct. Analysis of variance [12] of the correct-ratio (Table 2) shows a significant main effect of the angle between selectable viewpoints $\Delta\varphi$. There is no significant effect of the number of eyes on the correct ratio.

Source	SS	DF	MS	F	Sig
$\Delta\varphi$	6.805	4	1.701	8.303	.000
E	.605	1	.605	2.953	.086
$\Delta\varphi$ x E	.295	4	.074	.360	.837

Table 2: results of analysis of variance for correct-ratio

Figure 4 shows the effect of the variables on the correct-ratio and the 95% binomial confidence interval (CI) [3]. A one-sided t test [12] showed that the correct ratio for $\Delta\varphi=5.625^\circ$ is significantly higher ($p=0.05$) for the one-eye case than for the two-eye case.

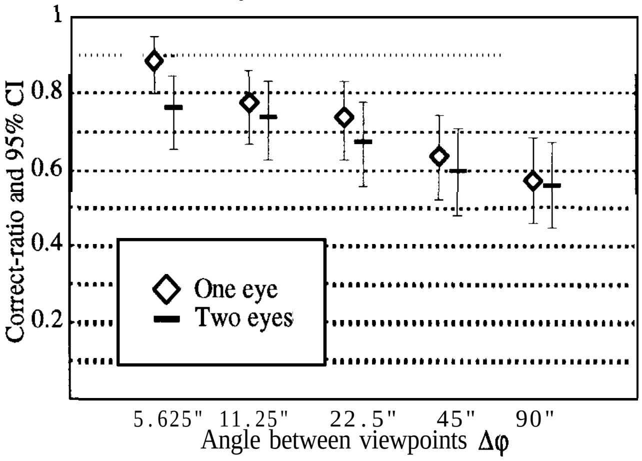

Figure 4: correct ratio and 95% CI

Analysis of variance of the response time (Table 3) shows a significant main effect of $\Delta\varphi$. Figure 5 shows the effect of the variables on the mean time and standard deviation.
An ad-hoc analysis of variance of the number of time-outs (a reaction time of more than 10 s.) showed a significant effect of $\Delta\varphi$ on it.

source	SS	DF	MS	F	Sig
$\Delta\varphi$	564,591	4	141.148	20.618	.000
E	1.069	1	1.069	.156	.693
$\Delta\varphi$ x E	20.622	4	5.155	.753	.556

Table 3: results of analysis of variance for T

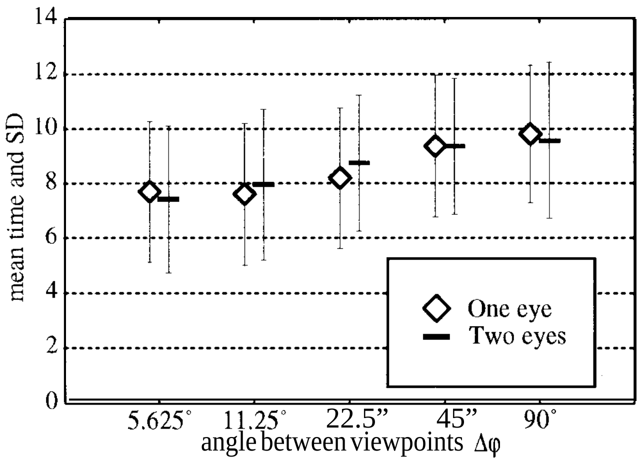

Figure 5: mean time and standard deviation in seconds

CONCLUSIONS

The DVWS works effectively, even with a small number of viewing angles. For the knot-solving task, the correct-ratio increases from about 0.55 to 0.9 as the angle between viewpoints decreases. A previous experiment [5,9] showed that for a simpler task even fewer viewpoints are sufficient.
The expected conflicts for the two-eye condition have no effect on the response time. If the angle between selectable viewpoints is 5.625° the correct-ratio is significant lower ($p=0.05$) when looking with two eyes in stead of one. For larger angles there is no significant difference.
The lower correct-ratio and longer response time of experiment [7] (see previous work) can be explained from the non-transparent wires and the low image resolution (less than 100x100) used there.
Concluding, the Delft Virtual Window System can be very useful for systems where only a limited number of viewpoints can be reached, made or calculated.

DISCUSSION, FUTURE RESEARCH

Some experiments suggest that fishtank VR is very sensitive for camera misplacements [8,11], while this and previous experiments showed that it poses no problems to the DVWS [5,9]. This may explain why fishtank VR usually presents a stereoscopic image, while the DVWS does not. Future work will examine the effect of the difference in projection methods between the fishtank VR and the DVWS and the effect of camera misplacements in more detail.

ACKNOWLEDGEMENTS

This work is supported by the Dutch Technology Foundation (STW) under grant DIO22-2732.

REFERENCES

1. Smets, G. J. F., Overbeeke, C. J. and Stratmann, M. H. Depth on a flat screen. Perceptual and Motor Skills 64 (1987), 1023 - 1034.

2. Overbeeke, C. J., Smets, G. J. F. and Stratmann, M. H. Depth on a flat screen: II. Perceptual and Motor Skills 65 (1987), 120.

3. Loosen, F. Statistische tabellen. Uitgeverij Acco, Leuven, Belgium, 1994.

4. Drascic, D. and Grodski, J. J. Using stereoscopic video for defence teleoperation: a Preliminary study. in Proc. SPIE 1915 Stereoscopic displays and applications IV (San Jose, CA, Feb. 1-2, 1993), pp. 58-69.

5. Pasman, W., Smets, G. J. F., and Stappers, P. J. Effects of image quality, number of selectable viewpoints and way to select the viewpoint in X-ray luggage inspection. submitted to Presence, 1995.

6. Sollenberger, R. L. and Milgram, P. A comparative study of rotational and stereoscopic computer graphic depth cues. in Proc. of the human factors society 35th annual meeting, (San Francisco, CA, Sep. 2-6, 1991). pp. 1452-1456.

7. Voorhorst, F., Overbeeke, C., and Smets, G. J. F. Movement parallax in endoscopy. Delft university of technology, Industrial design engineering, internal report, 1995.

8. Deering, M. High resolution virtual reality. Computer Graphics 26, 2 (July 1992), 195-202.

9. Pasman, W., Smets, G. J. F. and Stappers, P. J. Judging connections with a small number of viewpoints. Delft University of Technology, unpublished technical report, 1995.

10. Ware, C., Arthur, K. and Booth, K. S. Fish tank virtual reality. in Proc. INTERCHI '93. (Amsterdam, the Netherlands, April 24-29, 1993). uitgeverij IOS, pp. 37-42.

11. Arthur, K. W., Booth, K. S. and Ware, C. Evaluating 3D task performance for fish tank virtual worlds. ACM Transactions on Information Systems 11, 3 (1993), 239-265.

12. Hays, W. L. Statistics. CBS college publishing, New York, 1981.

13. Schwartz, A. Head tracking stereoscopic display. in Proc. of the SID 27, 2 (1986), Los Angelos, pp. 133-137.

14. Rekimoto, J. A vision-based head tracker for fish tank virtual reality: VR without gear. in Proc. of the VRAIS (Research Triangle Park, NC., 1995). IEEE Computer society press, Los Alamitos, CA, pp. 94-100.

15. Diamond, R., Wynn A., Thornsen, K. and Turner, J. Three dimensional perception for one-eyed guys or the use of dynamic parallax. In Computational Crystallography, David Sayre, Ed., Clarendon Press, Oxford, 1982, pp. 286-293.

16. Katayama, A., Tanaka, K., Oshino, T. and Tamura, H. A viewpoint dependent stereoscopic display using interpolation of multi-viewpoint images. in Proc. of the SPIE 2409 Stereoscopic displays and virtual reality systems II (San Jose, CA, 7-9 Feb. 1995). SPIE, Bellingham, 11-20.