INVESTIGATION OF SURFACE CHARACTERISTIC EFFECTS ON REAL-VIRTUAL OBJECT ALIGNMENT IN STEREOSCOPIC AUGMENTED REALITY

This paper reports a continuous virtual pointer alignment experiment carried out in a stereoscopic augmented reality environment. The object was to evaluate subjects' sensitivity to surface texture, target position at designated probe points on a cylinder real object surface, virtual pointer form and binocular disparity. The results confirmed the main findings from a previous experiment: that both surface texture and target position have significant influences. Subjective evaluation of virtual pointer form revealed that a three dimensional pointer is preferred over one and two-dimensional pointers. The contributions of size cue and resolution on binocular disparity effect are also discussed.

INTRODUCTION

Combining stereo computer graphics and stereo video images together within a unitary Augmented Reality (AR) display provides the potential for using stereoscopic images to solve a range of problems in hostile and inaccessible environments. These include robotic control in remote mining workplaces and estimating the size of anatomical structures in neurosurgery (Kim et al, 2000). To address such absolute measurement issues and provide guidance for AR interface design, a virtual pointer (computer generated stereo graphic cursor superimposed on a stereo video image) alignment task was conducted relative to a cylinder real object surface (Hou and Milgram, 2000).

In that earlier experiment, the effects of surface characteristics on alignment between graphic and real objects were investigated by interactively manipulating a *virtual stereographic pointer (VP)* relative to designated features in the stereo video image. It was found that surface texture and position on the target cylinder's surface both significantly affected placement accuracy. Those results suggested that depth cue conflicts plays a very important role in terms of surface texture effects; that is, highly textured surfaces facilitate higher localisation accuracy than low textured surfaces.

In general, whenever such a VP is placed in front of the surface of a real (video) object, i.e. the case in which binocular disparity cues and occlusion cues are consistent, subjects are able to shift attention easily back and forth from the pointer to the surface. On the other hand, whenever the virtual pointer is moved behind the surface, the two depth cues begin to conflict, because it should not be possible for an object to remain visible when placed behind an otherwise opaque intervening surface. Whenever that surface is sparsely textured, there are relatively few features to drive the stereoscopic fusion cue, so subjects are more easily able to reconcile the two conflicting cues and fuse both the real and virtual images. The result in such cases is that the real surface appears transparent; however, transitions through such surfaces are more gradual and are thus more difficult to detect. In other words, surface alignment performance is expected to be worse for cases in which surfaces appear transparent and transitions are gradual.

On the other hand, whenever the pointer is placed behind a highly textured surface, subjects are less able to overcome the tendency to fuse the surface texture features stereoscopically. In that case it is more difficult to reconcile the fact that the fused pointer is behind the fused surface yet still visible - a "perceptual impossibility", resulting in the breakdown of the VP into a double image. Although this cue conflict can bring some discomfort to perception such AR environments, an interesting practical question is whether this conflict can be used as an extra cue to assist AR interface design. In other words, our research question involves whether subjects will regard this breakdown phenomenon as a signal to judge proximity interactions between a virtual pointer and real video image objects and use this to detect the object surface?

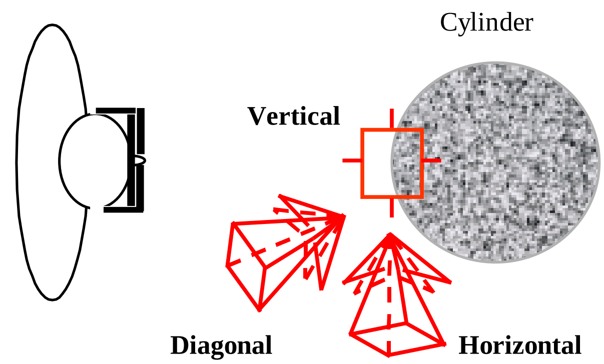
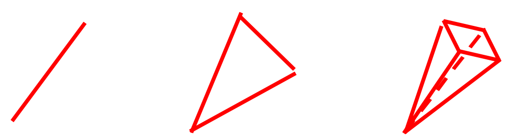


Figure 1. VP orientations in the earlier experiment (Top View)

Although this was also a goal of the earlier experiment, the results showed a positive bias (where the VP was typically placed in front of the real surface), with relatively large error ranging between 2 to 6cm. In reviewing the geometric design of the virtual pointer, it was noted that a possibility existed that the geometrical form of the pointer might have influenced the results. three different orientations of the VP were used: vertical, horizontal and diagonal, and each of those had four arms on the tip (see Figure 1). For the first two orientations, the pointer remained within a 2D plane tangential to the cylinder surface. The diagonal pointer was located within a plane that was at 45 degrees to both the tangent plane and the normal vector. Thus, there would be at least two third of the whole trials in which one of the arms touched and went behind the surface before the tip of the VP (which is the measuring point) touched. The cue conflict caused by the arm and the surface was perceived first, thus subjects might think that the VP already touched or went behind the surface without the tip touching the surface yet. We believe that this is the reason why the localisation error was big and the detected surfaces were in front of the real surface positions.

In order to test this hypothesis and thus to confirm if subject explore the surface by using the fusion breakdown as an extra cue, we took out the arms on the VP tip and redesigned its form (as an independent variable) for the present experiment. Based on the definition of stereoscopic cursor shape proposed by Barham and McAllister (Barham and McAllister, 1991), three types of VP with one, two, and three dimensions were designed, respectively (LINE, AREA, and VOLUME in Figure 2). With these three types of VP forms, the goal of the present experiment is to test surface texture, target position, VP form, and binocular disparity effects again, in addition, we also conducted the subjective evaluation on the VP form to investigate their preferences.



LINE **AREA** **VOLUME**

Figure 2. VP Form for the present experiment

METHOD

Similar to the earlier one, this experiment consisted of a psychophysical method of adjustment task involving the alignment of the VP with designated targets on the surface of the real object image, all of which were displayed using stereoscopic augmented reality. A 2x2x3x3 experimental design was used, comprising a combination of two textures (high and low density), two target positions (on the centre of the surface facing the observer and on the right side along the normal lateral plane), three VP forms (one, two, and three-dimensional), and three image disparities (crossed, 0, and uncrossed). The dependent variable measured was the error between the final VP placement and the actual position of the target on the surface of the real object.

Apparatus

Two target cylinders with 46cm diameter were used. Texture densities were different for two cylinder surfaces, and textures consisted of white dots randomly dispersed on a black background. The front surfaces of two cylinders were located 82cm, 92cm, and 102cm from the stereo cameras corresponding to *crossed* disparity, *no* disparity and *uncrossed* disparity, respectively. The stereo images of these cylinders were recorded and displayed on a SGI Indy workstation, and viewed through synchronised Imax liquid crystal shutter glasses by subjects at 48 cm from the screen.

Three VPs were computer generated arrows that appeared to hover within the stereo image upon which they were superimposed (only one VP for each surface image). They were diagonally oriented as the one in the earlier experiment, which was located within a plane that was at 45º to both the tangent plane and the normal vector. Three VPs were controlled with a Spaceball operating with only 3 translation degrees of freedom.

Procedure

The experimental task was to localise target points on the cylinder surfaces by manipulating the VP, for three VP forms, two textures, three binocular disparities and two target positions. Subjects used the Spaceball to move the VP in and out along X, Y, Z axes until it appeared to them to touch the surface of the object exactly at the designated target location. They then informed the experimenter that the alignment had been completed. Each experiment consisted of 6 randomised replications for each condition, for a total of 216 judgements. The experiment, including practice, took place over a span of three days, with each session lasting approximately two hours per day.

Paired Comparison

In order to obtain subjective assessment on different VP forms, a paired comparison was performed right after the whole trials were completed. Subjects were presented two interfaces side by side with different combinations of textured surfaces and VP forms. They were asked to manipulate the VP on both interfaces and give the ranking in terms of "ease of use" (which interface is easier to locate the surface), "transparency" (which surface looks more transparent – can be seen through when the VP is behind it), and "ease of fusion" (which combination is easier to fuse both surface and the VP). There were totally 15 randomised pairs of 6 combinations of VP forms and textured surfaces (see Table 1). Figure 3 illustrates one example of these pairs.

Table 1. Image Combinations for Paired Comparison

Image #	**1**	**2**	**3**	**4**	**5**	**6**
Texture Density	High	High	High	Low	Low	Low
VP Form	LINE	AREA	VOLUME	LINE	AREA	VOLUME

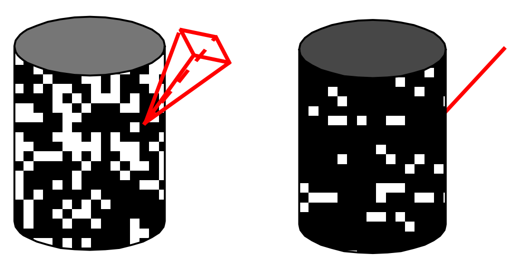


Figure 3. Paired Images for Comparison

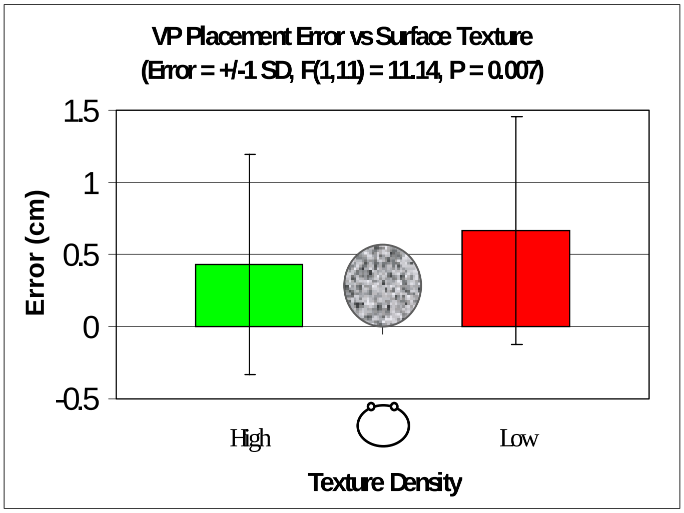


Figure 4. Effect of Surface Texture (High vs Low)

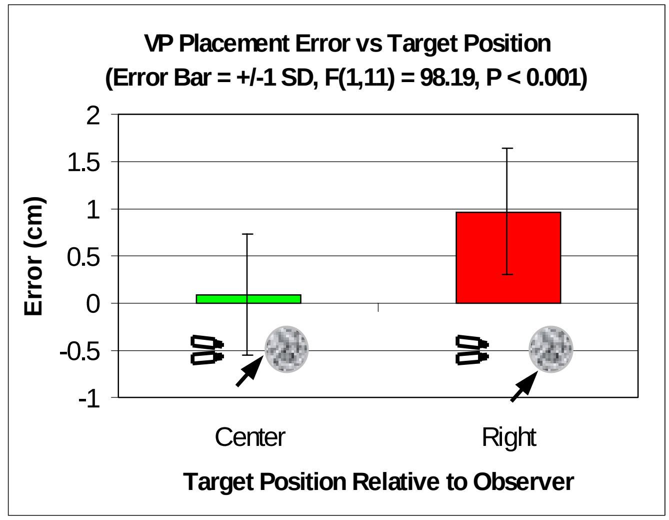


Figure 5. Effect of Target Position (Center vs Right)

Subjects

Twelve university students (7 male and 5 female) were screened using the RANDOT STEREOTESTS to participate the experiment. None of them knew about the design and the goal of the experiment. Where necessary, subjects wore appropriate optical correction.

RESULTS AND DISCUSSION

ANOVA was performed to analyse the experimental result. Figure 4 and 5 summarised the main experimental result. As hypothesised, surface texture and target position had significant effects on the placement accuracy ($F(1,11) = 11.14$, $P < 0.01$, and $F(1,11) = 98.19$, $P < 0.001$, respectively). This finding is consistent with what was found from the earlier experiment, in which highly textured surface had less placement error than low textured surface, and the performance was better when the VP was placed at a point along the centre of a surface relative to the observer's normal straight view angle (that is, looking straight at the surface) as compared to any other angle relative to the normal lateral plane. These figures also illustrated that the detected surface position is behind the real surface, and the placement errors were much smaller than those in the earlier experiment (less than 1cm). The post-experimental interview also supported the hypothesis that subjects were using the fusion breakdown (in their words: image "blurring" or "fuzzy") as a signal to locate where the surface is. Subjects pushed the VP into the surface until they saw a blurring or fuzzy image, then they started moving back the VP until the fusion difficulty disappeared (seeing clear image of the pointer and the surface) or much less than the previous position. Because it could be seen clearly when most part of the VP was outside (in front of) the surface (even though the tip was still inside -- transparency effect), they thought the VP was just right on the surface.

Although binocular disparity statistically had significant effect ($F(2,22) = 36.84$, $P < 0.001$) in the present experiment, it cannot be really claimed that disparity cue has significant impact on the probing performance because two other factors were brought in due to the specific experimental set-up. In order to keep the stereo cameras static thus to keep the same system precision and accuracy for three different disparities, instead of adjusting the sensitive stereo cameras, the target cylinders were moved to different positions corresponding to crossed, 0, and uncrossed disparities, respectively. Thus, the image for the crossed disparity (close to the observer) had biggest cylinder size and highest resolution, and the image for the uncrossed disparity had smallest cylinder size and lowest resolution. This is the reason why crossed disparity facilitated minimum error, and uncrossed disparity had maximum error in this case as Figure 6 shown.

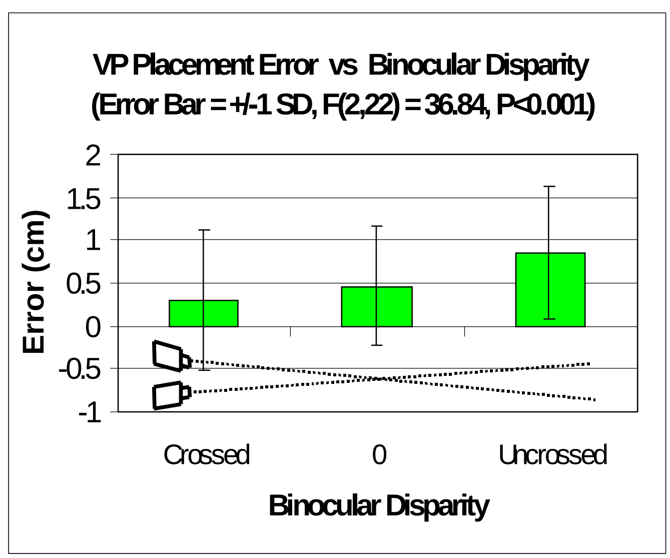


Figure 6. Effect of Binocular Disparity

Another prior hypothesis was that the VP form would also have significant effect, but the ANOVA analysis revealed no statistical significance ($F(2,22) = 1.655$, $P = 0.214$). However, the result from paired comparison did show subjects' preferences in terms of "ease of use", "transparency", and "ease of fusion" for 6 image combinations of VP forms and textures. Based on the proportion of choice of images in paired comparison made by 12 subjects, the probability of each choice was converted to Z score (Engen, 1971). Then mean Z scores were transformed linearly to illustrate psychological distance between 6 image combinations as shown in Table 2. It can be seen that subjects preferred low textured surface (image # 4, 5, and 6) rather than highly textured surface (image # 1, 2, and 3), because it is less likely to have fusion difficulty when the VP is placed behind the low textured surface (more transparent). They felt that it is easier to fuse both VP and surface and control (use) the VP for low textured surfaces, though they did not know that it is also easier to make bigger errors. However, mean Z scores for "ease of use" and "ease of fusion" indicated that subjects preferred three-dimensional VP (VOLUME) rather than one and two-dimensional VPs for both low and highly textured surface (image #3 and 6). VOLUME VP has more features along X, Y, and Z axes, it facilitates more depth perception in stereoscopic display than other two VPs. That is why subjects liked it the most. From the "transparency" rating, it can also be noticed that LINE VP is easier to facilitate transparency effect (for highly textured surface, Z score of image # 4 is bigger than #5 and 6; and for low textured surface, image #1 is bigger than #2). Overall, VOLUME one is the most favourable VP, and LINE is the least favourable one.

There are a number of perceptual issues related to Augmented Reality displays. This experiment re-examined the effects of surface characteristics, and confirmed not only surface texture and target position effects on the alignment between real and graphic objects, but also testified the existence of depth cue conflict between binocular disparity and occlusion on the fundamental level. In addition, discovered this conflict as an extra cue to detect the interaction between virtual and real objects in stereoscopic AR environment. Coupled with subjective evaluation on the VP form, these findings will definitely facilitate the design of more efficient probing tools in AR applications. The immediate practical application of this study is to use a VP effectively for guiding a drill relative to a variety of mine faces in a remote mining task. In order to increase our understanding of human perception of curved surfaces and thus provide more guidelines for AR interface design, further investigations are needed, such as the existence of optimal direction to approach the normal of a curved surface and how sensitive our visual perception to surface texture will be.

REFERENCES

1. Barham, P. T. and McAllister, D. F. "A Comparison of Stereoscopic Cursors for the Interactive Manipulation of B-Splines", *Proc. SPIE 1457 – Stereoscopic Displays and Applications II*, pp 18-26, 1991.
2. Engen, T. "Psychophysics: II. Scaling Methods". Chapter 3 in Kling, J. W. and Riggs, L. A.(ed's), *Woodworth & Schlosberg's Experimental Psychology*, Third Edition, Holt, Rinehart & Winston, pp 47-86, 1971.
3. Hou, M. and Milgram, P. "Effect of Surface Characteristics on Alignment of Graphic and Real Objects in A Stereoscopic Augmented Reality Environment". *Proceedings of the IEA2000/HFES2000 Congress*, Vol. 3, pp476-479, 2000.
4. Kim, M.Y., Drake, J.M. and Milgram, P. "Virtual tape measure for the operating microscope: System specifications and performance evaluation". *Computer Aided Surgery* 5:148-155 (2000).

Table 2. Paired Comparison Results

	Linearly Transformed Mean Z Score					
Ease of Use	0	1.15	1.71		2.26	3.16
Image #	1	2	3	4	5	6
Transparency	0	0.25	0.89	2.97	3.82	4.37
Image #	2	1	3	6	5	4
Ease of Fusion	0	0.69	0.73	2.08	2.14	2.26
Image #	1	2	3	5	4	6