Perception and Mental Rotation of 3D-Freeform Surfaces in an Immersive Projection System

Andrea Gaggioli, Ralf Breining

Competence Center Virtual Reality
Fraunhofer Institute for Industrial Engineering,
Nobelstr. 12, 70569 Stuttgart, Germany
ralf.breining@iao.fhg.de

Abstract. This paper reports on three experiments performed to examine the effects of different visualization techniques of 3D computer-generated freeform surfaces on subjects' perceptual and cognitive performance while doing CAD-related activities in an immersive VR system. Experimental perceptual and cognitive tasks included depth size estimation of a single 3D object (exp. 1), estimation of depth differences between two 3D objects (exp. 2) and mental rotation of 3D objects (exp. 3). Dependent variables were accuracy (exp. 1, 2 and 3) and response time (exp. 3). The visualization techniques we investigated were presence versus absence of binocular disparity, four different types of graphic image (wireframe, flat shading, Gouraud shading and Gouraud shading with surface normals) and two levels of shape complexity.
Results showed a positive effect of binocular disparity on perceptual performance (esp. 1), in particular when concave 3D shapes were used as stimuli (esp. 2), but a limited positive effect of stereopsis on mental rotation. Furthermore, results indicated that subjects were faster in mentally rotating 3D shapes rendered with more realistic techniques, whereas perceptual estimates were found more accurate and easier when observers were presented with less realistic rendered surfaces (exp. 1-2).

1 Introduction

A major use of 3D computer graphics is in design process, particularly for engineering and architectural systems, but almost all products are now computer designed. In the next future, CAD methods will be extensively employed in conjunction with 3D immersive displays, which may dramatically improve the possibilities of visualization and interaction offered by common 2D display CAD workstations. Thus, careful evaluation needs to be made as to how the computer-generated object is represented on 3D immersive display during the design process. It is commonly recognized that a computer-generated 3D object should be a) an accurate description of the model being designed b) presented in a realistic and integrated format, so that it can be visualized and interpreted without introducing uncertainty regarding to the represented proprieties [2]. To ensure that the displayed 3D image will satisfy these requirements, it is important focusing on which perceptual/cognitive operation is performed by the user during the design process. The main purpose of this research was to investigate the effects of stereo vision and monocular cuing techniques on subjects perceptual and cognitive abilities that are relevant in CAD-related activities, such depth estimation and mental rotation of computer-generated freeform surfaces.

Depth can be artificially simulated by presenting on a display's two-dimensional surface properties of the external world that results in the sensation of depth, called depth cues. Depth cues can be classified as monocular or binocular. Retinal disparity is an important binocular depth cue and is caused by the fact that each of our eyes sees the world from a different point of view. These two different perspectives of the same image are then combined by the brain into a single image by means of a process called *fusion*. The resultant sense of depth is called *stereopsis*. Monocular depth cues include light and shade, relative size, interposition, textural gradient, motion parallax and perspective. For some design (CAD) applications, objects are first displayed in a wireframe outline form that shows the overall shape and internal features of objects (i.e., the interior of a vehicle). Wireframe displays are useful to perform animations because the calculations for each segment of the animation can be performed more quickly when rendered surfaces of one object are not displayed. When object designs are complete, or nearly complete, realistic lightning models and surface rendering are applied to produce displays that will show the appearance of the final product [2]. Shading is used widely to depict surface structure and depends on variation of reflected light intensities according to the relationship between surface orientation and illumination direction. It relies on the ability of the human visual system to abstract shape from a physical or simulated scene illuminated from one or more light source.

Although it is known that monocular cues and steropsis can contribute individually, or in combination, to an appreciation of the object three-dimensional shape, there is little information on the relative merits of them. This is important in the case of virtual objects since incorporation of some visual cues is associated with a significant computational expense. For instance, stereoscopic vision requires two display channels to be rendered in a frame interval. Thus, an understanding of the relative benefits of these monocular and binocular coding techniques helps in judging the overall value of including a particular cue. In so doing, virtual reality application can be tuned and optimised for real time performance.

2 Experiment I: Depth Estimation of a 3D Object

The main purpose of the first experiment was to determine how stereo vision and different monocular coding techniques affect the ability to estimating depth of a 3D computer-generated object. Previous research on human distance and depth perception in 3D environments has addressed two basic questions, that is, which visual cues provide effective depth information and how does the effectiveness of given depth cues changes as a function of the viewing distance. Most studies have focused on subjects ability to estimate egocentric distance and relative distance, while there exists poor research about how depth of single objects is perceived. Real objects have depth and are located phenomenally as well as physically at different distances [3]. That is, objects have an egocentric distance (i.e., the distance of the object from an observer) and a relative distance (i.e., the distance of objects from each other). Objects also have depth in that they are perceived as three-dimensional and some parts of an object look farther away than do other parts [6]. This type of depth perception can be described as the ability of perceive variations in egocentric distance when looking at different points located on a 3D object. Stereo vision and four monocular coding techniques were investigated. Monocular coding techniques were wireframe (with hidden edges removed), flat shading, Gouraud shading [4] and Gouraud shading with surface normals.

2.1 Method

Subjects. Twenty-four subjects served as participants in the study (mean age = 29.4 years). All subjects had normal or corrected-to-normal vision and reported no experience with virtual reality, no experience with CAD software and upper-intermediate experience with computer.

Design. The experiment consisted of two types of object geometry (simple versus complex), four types of graphic images (wireframe, flat shading, Gouraud shading and Gouraud shading with surface normals) and six depth values (15, 30, 45, 60, 75, 90 cm) as completely crossed factors with stereopsis (present or absent). The serial order of depth values was determined by a Latin square arrangement and was the same for all participants. All factors were within subjects.

Apparatus and Stimuli. 3D images were created using these software 3D Studio Max, release 2.5, and displayed on the front wall of a four-walls CAVE capable of both stereoscopic and monoscopic modes. The screen is 300 cm wide and 300 cm high with a resolution of 960 x 960. Stereoscopic condition was created using CrystalEyes time multiplexed LCD shutter glasses synchronised to the display monitor. Two basic 3D free-form shapes were created (see Fig.1 and Fig.2). The simple one (a) was represented by a medusa-like object, the complex shape (b) was a terrain. Both objects were 50 cm wide. A squared base was added to the scene because during a pilot-experiment subjects reported that stimuli "floated" above the projection plane.


Fig. 1


Fig. 2

Fig. 1 - 2. Stimuli used in exp. 1 (both objects are here rendered with Gouraud shading)

Procedure. Participants were told the purpose of the research and given specific instructions regarding the estimation task. In estimating depths, participants were required to report their estimates in centimetres. In order to facilitate the task, subjects were informed about their egocentric distance from the projection plane and given the edge of the square (100 cm) that represented the base of the 3D objects. Following practice, they estimated depth on each of 48 trials (6 trials for each rendering method; 24 trials for each geometry type) with stereopsis (stereo condition), took five minutes break, then estimated depth to another identical set of 48 trials without stereopsis (mono condition). From the participant's perspective, the depth value that the displayed object had on any given trial was random. In order to control the possible effect of sequence for the rendering conditions, four different sequences were arranged using a 4 x 4 latin square.

2.2 Data analysis and results

In order to analyse data, two ANOVAs were performed. The first used depth-difference estimates as the dependent measure and the second used relative errors as the dependent measure for accuracy. Relative error was calculated as follows: Relative error = (Depth estimate - True depth) / True depth.

This represents the percent error in an estimate relative to the true depth difference, with the sign indicating the direction of the error (when the sign is negative, it indicates underestimation; when the sign is positive, it indicates overestimation). Stereopsis, type of geometry, type of rendering and size of depth-difference were the independent variables in both analyses. Fig. 3 illustrates the means of the estimates in stereo and mono condition. The line with crosses represents perfect performance.

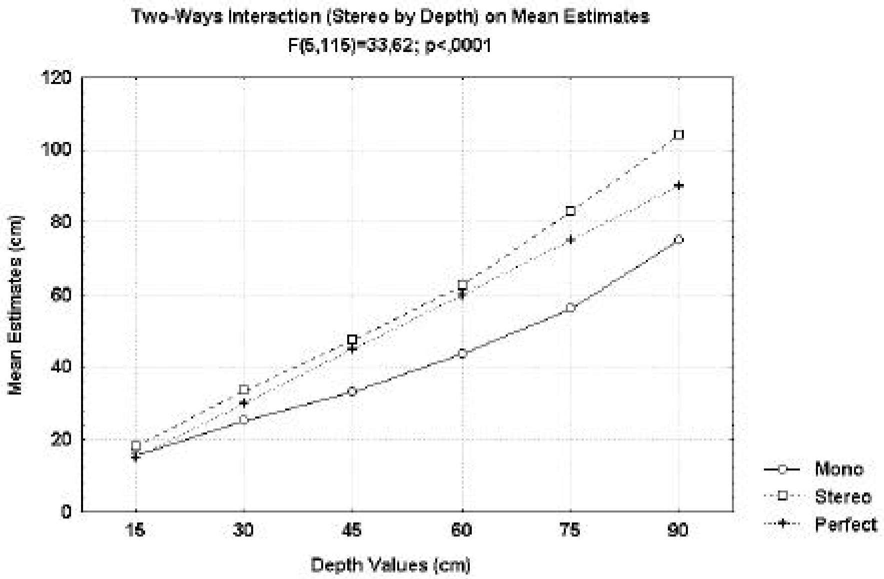

Figure 3. Effect of Stereopsis and Depth Values on estimates (exp. 1).

Fig. 3 shows that observers generally underestimated depth of 3D objects presented in mono, while performance in stereo was clearly more accurate. This is confirmed by the high significant effect of stereopsis in the ANOVA performed on relative error

(F(1,23)=69,65; p < 0,000001). The effect of object complexity on accuracy is reported in Fig. 4. This graph shows that estimates were far more accurate for the simple shape than for the complex shape.

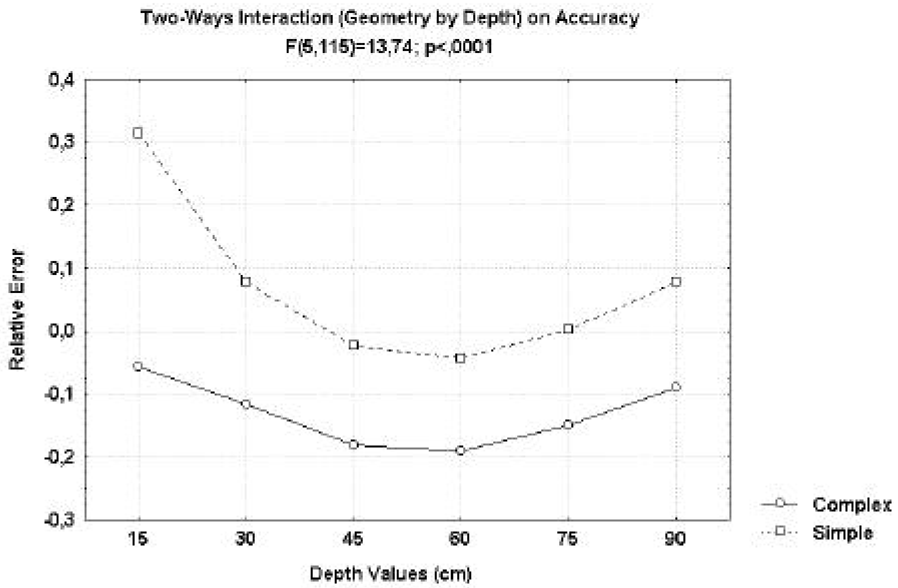

Figure 4. Effect of object complexity on relative error.

Fig.5 illustrates the effect of graphic coding techniques on accuracy. The main effect is very significant. A Tukey's HSD test was performed to assess pair wise comparisons. Results are summarized in Table 1. It was found that mean relative errors for Gouraud shading differed significantly from each other rendering method, and that there was a significant difference among Wirefame and Normals. Normals and Flat Shading determined more accurate estimates than Wireframe and Gouraud shading and the worst performance was determined by Gouraud shading.

MCT	Wireframe	Normals	Flat Shading	Gouraud Shading
Wireframe		*0,013578*	0,467195	*0,011390*
Normals	*0,013578*		0,349044	*0,000151*
Flat Shading	0,467195	0,349083		*0,000230*
Gouraud Shading	*0,011390*	*0,000151*	*0,000230*	

Note: Values indicate probability, and italic values reflect significant pair wise comparisons.

Table 1. Post hoc analysis of monocular coding techniques (dependent variable = relative error)

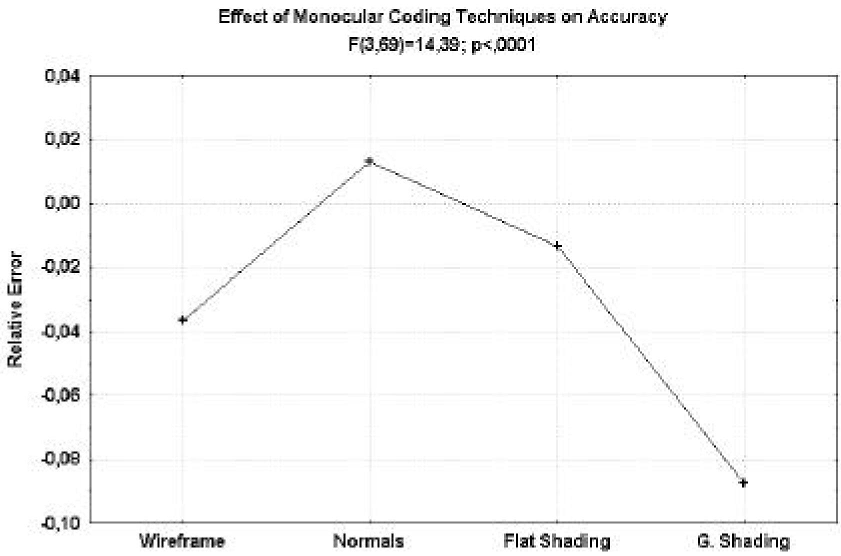

Figure 5. Effect of monocular coding techniques on relative error.

2.3 Discussion

Estimates and accuracy were significant affected by depth values, stereopsis, object complexity and rendering methods. The highest relative error was determined by the smallest depth value (15 cm). A post-hoc analysis performed on the interaction geometry type by object depth revealed that this effect was principally due to overestimation of the smallest depth value of the simple object ($p < 0{,}01$).

Subjects estimated depth of 3D objects more accurately when objects were displayed with binocular disparity than when only monocular cues were provided. Thus, retinal disparity confirms to be a very important depth cue, that provides effective perceptual information not only in estimating egocentric/relative distances but also in estimating distances between parts of an object that lay in different depth positions.

The effect of object complexity was highly significant. Complex objects depth was underestimated while simple objects depth - with the exception of the depth value 15 cm - was estimated more accurately. This difference could be due to the fact that irregular changes in the orientation of the surface of the complex object (a terrain-like 3D shape) did not allow subjects to detect the exact position in space of the top of the object.

Another interesting result of this study is that standard CAD monocular coding techniques affect depth estimates in different ways. Depth of wireframe objects was estimated more accurately than depth of Gouraud shaded objects but this result was inverted when normals were added to smooth shaded objects. This finding can be interpreted through results by Koenderink et al. [5] who showed that most of the relief of a shape is determined by visual contour and that shading adds very little to the "solidity" of an object if other cues are not available. Todd and Mingolla [10] performed three experiments to examine the perceptual salience of shading, texture, specular highlights and directions of light sources in providing information about the

3D structure of a cylinder. The results indicated that the shininess of the surface enhanced the perception of the curvature, but had no effect on the perceived direction of illumination, and that shading was generally less effective than the gradient produced by the texture for depicting surface in three dimensions. In our study, texture was not a surface feature but the addition of small normals across smooth-shaded surfaces appeared to be incremental to accuracy. This enhancement may depend on the fact that normals produced a gradient which served as additional depth cue. Another explanation could be that normals orientation allowed subjects to draw more detailed inferences concerning local attitude (slant and tilt) of the 3D surface. In order to disambiguate between these two explanations, however, further research is needed.

3 Experiment II: Estimation of Depth Differences Between 3D Objects

The main purpose of the second experiment was to investigate how stereo vision and different monocular coding techniques affect the ability to estimating depth differences between two 3D computer-generated forms displayed simultaneously. This task was supposed to be perceptually and cognitively more demanding than the task of Exp. 1, because subjects were to estimate depth of both objects and then to calculate their difference. Investigated factors were stereo vision, the four monocular coding techniques described in the previous experiment and type of object geometry (concave versus convex).

3.1 Method

Apparatus and Stimuli. The apparatus was the same used in experiment 1. Twelve couples of convex spheroids (Fig. 6a) and twelve couples of concave spheroids (Fig. 6b) were designed using the software 3D-Studio Max. For each objects group, the direction of depth difference was balanced: for the first six objects, the spheroid positioned on the right side of the square was greater in depth than the spheroid positioned on the left side, while for the remaining six objects the order was inverted.

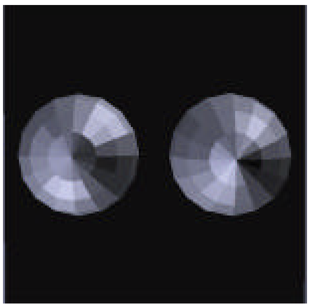
Fig. 6a

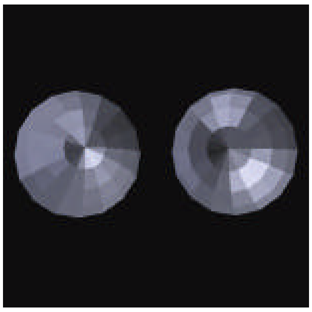
Fig. 6b

Figure 6a-b. Stimuli used in exp. 2 (both objects are here rendered with flat shading)

3.2 Data analysis and results

In order to analyse data, two ANOVA were performed. The first used depth-difference estimates as the dependent measure and the second used relative errors as the dependent measure for accuracy.

Fig. 7 illustrates the means of the estimates in Stereo and in the Mono condition. The line with crosses represents perfect performance. Fig. 7 shows that although observers generally underestimated differences in depth between the two objects, the level of underestimation was significantly reduced when binocular disparity was provided: actually, estimates in Stereo typically averaged 90% of the true depth difference, while estimates in Mono averaged 71% of the true depth difference.

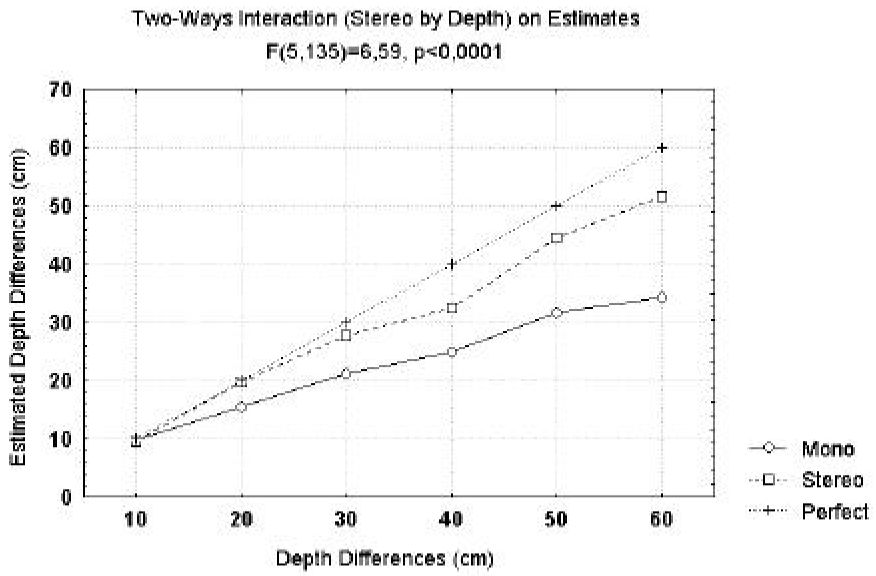

Figure 7. Effect of Stereopsis and Depth Values on estimates.

3.3 Discussion

Subjects generally underestimated differences in depth between the two three-dimensional shapes. Corroborating the results of the first experiment, the errors in depth estimation were found greater when objects were displayed without stereopsis. This confirms that stereopsis represents a powerful depth cue when the task is perceptually demanding [11]. The interaction between stereopsis and geometry type on relative error is also of particular interest. Depth differences for concave and convex objects were far underestimated in mono, and the underestimation for concave objects was found significantly greater than the underestimation for convex objects. When stereo vision was provided estimates for concave objects became as accurate as estimates for convex objects. This finding supports the assumption that the advantage of this depth cue becomes more pronounced when monocular information are ambiguous or insufficient in depicting the shape of a three-dimensional object [7].

Actually, some subjects reported difficulties in perceiving concavity without ambiguity when only monocular cues were provided, especially if concave shapes were rendered using the Gouraud shading technique. This subjective impression is corroborated by mean results for type of graphic image, that reveals a trend of underestimation of depth differences for Gouraud shaded objects, although the magnitude of the error was less significant than in Exp. 1.

4 Experiment III: Mental Rotation of 3D Objects

The previous experiments have provided evidences that participants' accuracy in estimating depth of 3D virtual objects with variable geometry is affected both by stereopsis and monocular coding techniques. In particular, it was found that the absence of binocular disparity led to an impoverished performance, whereas – more surprisingly – the use of less realistic types of graphic image (i.e. wireframe and flat shading) and the addition of surface normals improved accuracy in estimating depth. The first objective of the third experiment was to assess whether the same pattern of results could be obtained with subjects performing a more complex cognitive task, the mental rotation of three-dimensional objects. This standard cognitive task consists of presenting subjects with pairs of drawings of stationary 3D block objects and then asking participants to determine whether the presented images are different objects or different angular orientations of the same object (Shepard, 1971). In order to answer this questions, subjects mentally rotate one of the two objects into congruence with the second object for comparison. This assumption is based on the observation that the time required to make a correct judgment of "same" increase linearly with the angular difference between the objects.
The mental rotation paradigm has been applied in the computer graphics research because it represents one of the few available objective methods to assess user's cognitive performance under different visualization parameters of the computer-generated object.
The main goal of the third experiment was to further investigate the effects of stereopsis and monocular coding techniques on the cognitive manipulation of 3D structures. In particular, we were interested to assess whether the positive effect of binocular disparity observed in exp. 1-2 is strictly related to the perceptual task at hand (depth estimation), or is extended to more complicated cognitive tasks (i.e. mental rotation).

4.1 Method

Design. The experiment consisted of two levels of object geometry (simple, complex), four types of graphic images (wireframe, flat shading, Gouraud shading, Gouraud shading with surface normals) and six angles of relative rotation (0°, 36°, 72°, 108°, 144°, 180°) as completely crossed factors with stereopsis (present or absent). Each of these combinations of factors were presented to subjects in two pairs of slides. One of these pairs was a "same" pair while the other was a mirror image or "different" pair. All factors were within subjects. The 96 treatment combinations were presented to subjects in random order.

Apparatus and Stimuli. The apparatus was the same used in experiments 1-2, except that subjects were given two buttons to indicate their responses. The 3D computer generated objects (see Fig. 8) were modelled after the 3D figures portrayed in the experiment of Barfield (1988).

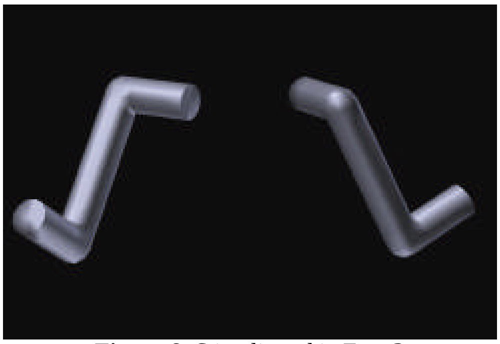

Figure 8. Stimuli used in Exp. 3

4.2 Data analysis and results

Response Time. Results of the ANOVA performed on response time data indicated a significant main effect for type of graphic image ($F(3,51) = 6,36$; $p<0,001$). A post-hoc contrast revealed that the wireframe images produced significantly slower response times than Gouraud shaded images. The ANOVA results indicated also significant main effect for angle of deviation ($F(5,85) = 13,85$; $p<0,0001$) on response time. Fig. 9 shows that response time increases almost linearly with angle of rotation, suggesting that subjects mentally rotate 3D objects. There was also a significant main effect of relative orientation, with mirror objects producing slower response time than "same" objects ($F(1,17) = 19,53$; $p<0,0005$). The effects of stereopsis and object complexity were not significant.

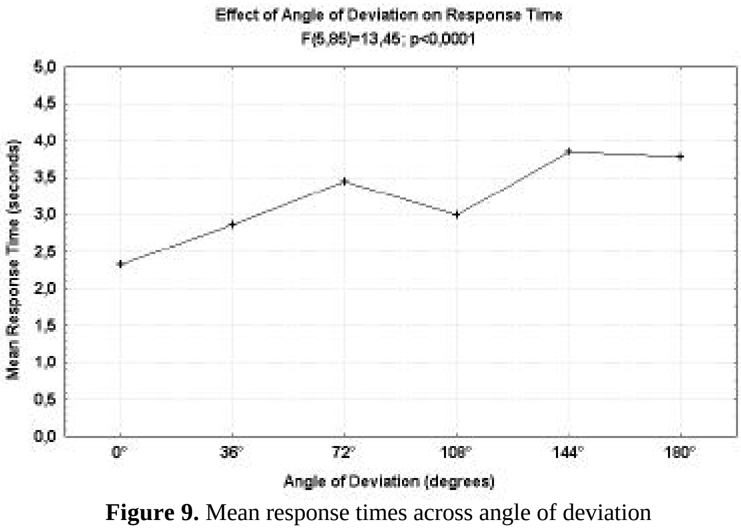

Figure 9. Mean response times across angle of deviation

Response accuracy. The overall error rate for the experiment (defined as the percentage of trials in which subjects responded incorrectly) was 7,2 %. This compares good with results from previous studies (5,9 %: Barfield [1]; 6,1%: Yuille and Steiger [12]; 3,2%: Shepard [9]). The ANOVA performed on percentage of error revealed significant main effects for stereopsis ($F(1,15) = 5,88$; $p<0,05$) with subjects more accurate in the stereoscopic condition, and for angle of deviation ($F(5,75) = 2,37$; $p<0,05$). In particular, it was found that the angle 180° was the most difficult rotation angle in the experiment, with a error rate of 10 %. The effects of the remaining factors were not significant.

4.3 Discussion

The main objectives of the third experiment were a) to assess whether providing the cue of stereopsis would help subjects to better perform a mental rotation task and b) to verify whether different monocular coding techniques of a 3D computer generated object impact performance differently. Results of this experiment indicated that the stereopsis cue helped subjects perform the mental rotation task, although this effect was limited to the accuracy with which participants discriminated 3D objects.

As concerns the effect of monocular coding techniques, results indicate that the response time was significantly slower when subjects were presented with wireframe images. This finding is in agreement with results of previous studies [1; 2] that showed that the discrimination task is better when realism cues are added to the object surface and in particular when shading procedures are used. Thus, the present study confirms that the use of advanced rendering can not only enhance the aesthetics and the form faithfulness of a computer generated object (as shown by subjective judgment studies, e.g. Sanford, 1987 [8]), but it can also improve the user's ability to visualize and mentally manipulate the structure of such object.

5. General Discussion

The most convincing conclusion suggested by the results of this research is that the stereopsis cue is useful and sometimes even necessary during CAD-related activities. As shown by experiments 1-2, subjects estimated depth of 3D objects far more accurately when objects were displayed with binocular disparity that when only monocular cues were provided.

As concerns the relative merits of monocular coding techniques, the three experiments performed produced mixed results. In particular, in exp. 1-2 it was found that the use of less realistic types of graphic image (i.e. wireframe and flat shading) and the addition of surface normals significantly improved subjects' accuracy in estimating depth; on the other hand, the third experiment indicated that the use of advanced rendering (Gouraud shading) enhanced the user's ability to visualize and mentally manipulate the structure of 3D objects. One way to account for these differences is to postulate that although both tasks imply the mental representation of the three-dimensional structure of the object, they involve different perceptual/cognitive processes. This could explain why the manipulation of specific monocular coding techniques (e.g. Gouraud shading) can determine different effects depending on which task is performed by subjects. If this assumption is correct, it would be cost-effective to give the user the possibility to customize the visualization method according to which CAD activity he/she is actually carrying out. In example, when the user is performing tasks that imply the direct estimation of the 3D object size, he could select the wireframe modus, which also has the advantage to be associated with small computational expense. Surface rendering could be applied to produce displays that assist the users in doing more cognitively demanding activities, such as rotating a figure to perform a visual interference check or to evaluate the overall aesthetic qualities of a digital mock-up.

References

1. Barfield, W., Sanford, J., Foley, J. (1988). The mental rotation and perceived realism of computer-generated three-dimensional images. *International Journal of Man-Machine Studies*, 29, p. 669-684.
2. Brown, M.E., Gallimore, J.J. (1995). Visualization of Three-Dimensional Structure During Computer-Aided Design. *International Journal of Human-Computer Interaction*, 7(1). p. 37-56.
3. Dember, W.N., Warm, J.S. (1979). *Psychology of perception* (2nd ed.). New York: Holt, Rinehart & Winston.
4. Foley, J., Van Dam, A. (1982). *Fundamentals of Interactive Computer Graphics*. Reading, MA: Addison-Wesley.
5. Koenderink, J.J., Van Doorn, A.J., Christou, C., Lappin, J.S. (1996). Shape constancy in pictorial relief. *Perception*, 25, p. 155-164.
6. Matlin, M.W., Foley, H.J. (1992). *Sensation and Perception* (3rd ed.). Boston: Allyn and Bacon.
7. Pepper, R. L., Patterson, R. (1986). Stereo TV improves operator performance under degraded visibility conditions. *Optical Engineering*, 20, p. 579-585.
8. Sanford, J., Barfield, W., Foley, J. (1987). Empirical studies of interactive computer graphics: Perceptual and cognitive issues. In *Proceedings of the Human Factors Society 31st Annual Meetting Volume I*, p. 519-523.
9. Shepard, RN., Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, p. 701-703.
10. Todd, J.T., Mingolla, E. (1983). Perception of Surface Curvature and Direction of Illumination From Patterns of Shading. *Journal of Experimental Psychology: Human Perception and Performance*, 9 (4), p. 583-595.
11. Yeh, Y.Y., and Silverstein, L.D. (1992). Spatial judgments and stereoscopic presentation of perspective displays. *Human Factors*, 34, p. 583-600.
12. Yuille, J.C., Steiger, J.H. (1982). Nonholistic processing in mental rotation: some suggestive evidence. *Perception and Psychophysics*, 31, p. 201-209.