International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor (2018): 7.426

Sampling Static Perspective Cues

Ajitkumar Sharma[1], Dr. Kiran Kumari[2]

[1]Department of BCA, B.R.A. Bihar University, Muzaffarpur, Hajipur, Vaishali, Bihar, India-844101
[2]Department of Physics, B.R.A. Bihar University, Muzaffarpur, Hajipur, Vaishali, Bihar, India-844101

Abstract: ***This paper primarily deals with sampling artefacts in interactive, immersive 3D CGI. However, before we can consider the perception of sampled 3D images in these dynamic situations, we need to analyse the presentation and perception of sampling in static images. First, we use the static case to establish some of the assumptions and conventions we use in this thesis to describe perspective geometry. Second, we identify and discuss the effects of spatial sampling artefacts in the presentation of perspective depth. Then we describe experiments performed to identify the visual contexts in which these artefacts are likely to impair the judgment of depth. Finally, we describe and analyse two effective methods for ameliorating these artefacts. Given a vanishing point located at the centre of the screen and the number of pixels, (n_h, n_v), we can compute the exact location of the 3D point in 2D screen coordinates: $x_s = (x - e_x) \cdot n \frac{h}{} \cdot e^{z} y_s = (y - e_y) \cdot {}^{nv} \cdot {}^{ez} s_h e_z - z \quad s_v e_z - z$ To get the sampled location of a point, we round (x_s, y_s) to the centre of the nearest pixel. In the equations above, we assume the line of sight orthogonally intersects the centre of the screen, a useful and frequently used convention. However, the location and orientation of the viewpoint plays a critical role in the perception of scene layout. It is impossible for a user to discriminate two points separated along the line of sight because the points share the same location on the screen. However, if the viewpoint is rotated around the two objects so their difference lies perpendicular to the viewer's line of sight, discriminating the two points is substantially easier.***

Keywords: Sampling Position, Sampling Size, Perceptual Implications, Methodology, Effect of Separation on Location and Size Judgments, Detectability of Sampled Perspective Cues

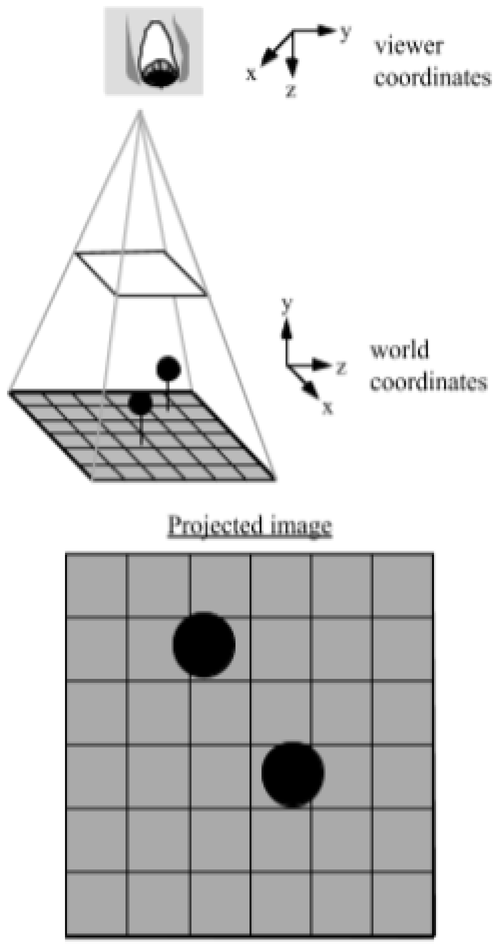
1. Introduction

Over the past forty years, advances in display and computing technology have revolutionized the interface between man and machine. Now that people can interact with rich, realistic, 3D graphics with relatively low cost equipment, the time has come to focus on designing our systems so that we maximize their capabilities in the ways most effective for the user.

Man-machine interfaces found in simulator and teleoperation systems have laid the groundwork for completely computer-generated or virtual environments. *Simulators* are training systems that display computer-generated scenes based on real-world situations. *Teleoperation systems* extend a person's ability to sense and manipulate the world to a remote location. The control and display devices in these systems are often computer-controlled. *Virtual environments (VEs)* are computer generated experiences that may seem real but are not required to match any of the rules of the real world.

Changing viewpoint can increase the amount of information provided about the separation of the two objects. The left column shows a typical VDS viewing situation and the projected image those results. The right column shows how changing the orientation of the scene relative to the viewer increases the distance between the objects depth in the projected image.

In head-tracked systems, we cannot select just one geometric viewpoint since the real viewpoint is tied to the user's head position. This head-coupled viewpoint is considered a major component of immersive VDSs ([1]). In these systems, the user has the ability to choose the best viewpoint for viewing a scene.



In desktop VDS, however, the control of the geometric viewpoint is a significantly more difficult matter. While immersive displays use head tracking as an intuitive interface to geometric viewpoint control, some other method must be used for choosing a geometric viewpoint in desktop VDSs ([2]). Typically, the location of the geometric viewpoint is chosen so that the real FOV matches the geometric FOV. However, the orientation of the scene relative to the viewer can be changed easily. In this thesis, we assume the user is looking at the centre of the screen and the line of sight is perpendicular to the VDS surface, as in Figure 5.1. We assume that the GFOV matches the real FOV, regardless of the orientation of the scene relative to the viewer.

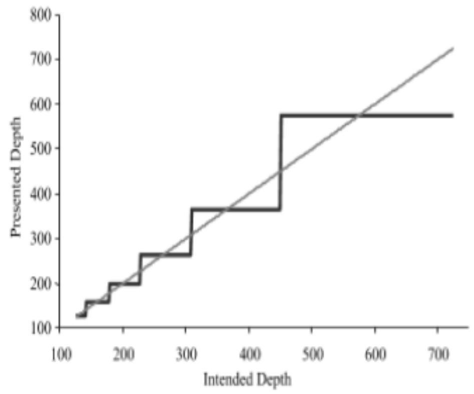
The main assumptions used throughout this thesis are:

- The orientation and location of the users real viewpoint is static
- The real location of the viewpoint matches the geometric location of the viewpoint
- Geometric and real FOVs are equivalent
- The viewer is looking at the centre of the screen
- The VDS surface is flat
- Pixels are square
- Pixels behave uniformly across the VDS surface

Sampling Position:
Positional inaccuracies due to sampling are relatively simple to compute. In 2D, location is rounded to the nearest pixel, resulting in a maximum error of half a pixel in either the vertical or the horizontal dimension. Similarly, in 3D, the positional error in projected points is at most a half-pixel. However, error in projected location may represent a significant error in depth. At close distances, a half-pixel error in the projected vertical or horizontal position may only represent a small inaccuracy in location,

Presented depth as a function of intended depth. The black line shows the sampled depth; the grey line shows the unsampled depth.

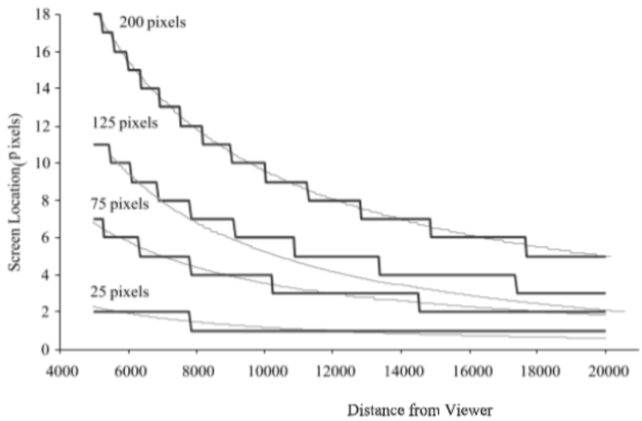
whereas an object far away suffers from significantly more error. Figure 5.3 shows the corresponding amount of error in distance caused by an error of plus or minus half a pixel in projected location.



At first glance, the decreased accuracy in the representation of depth that occurs at greater distances may seem to mirror the HVS decrease in spatial acuity with distance. However, spatial acuity actually improves as the target recedes in distance up to 5-10m. Furthermore, spatial acuity measured at one distance is a poor predictor for acuity at another distance. Therefore, the inaccuracy in depth due to sampling results in decreased performance that does not match the behavior of the HVS in the real world.

For example, the ability to compare two points in depth is significantly hindered in VDS with limited spatial resolution. If the two points are far from the viewpoint, they appear to be at the same position, although they may be separated by a large amount. For low-resolution VDSs, this problem is exacerbated. The steps in depth are fewer and larger, and depth acuity suffers accordingly.

From the viewpoint assumed, we know that the vanishing point is in the centre of the screen; therefore, points separated only in depth along the line of sight are indistinguishable. The distance of a point from the line of sight determines how much its projection changes as a function of distance from the viewer.



Sampled screen location as a function of distance from the viewer (along the line of sight). Different lines represent different distances from the line of sight. Grey lines show the unrounded values.
When a projected point is less than half a pixel from the projected vanishing point, further increasing depth has no perceivable effect. This occurs at a depth, z_v, which is the lesser of the horizontal and vertical vanishing distances, (z_x, z_y):

$$z_v = \cdots zz_{xy} \quad \text{ifotherwise} \quad z_x \geq z_y$$

Sampling Size:
When we consider lines and polygons parallel to the VDS surface, the lack of accuracy caused by using linear perspective to represent a point in depth results in additional artefacts. The number of pixel steps that occur as a point recedes in depth are a function of its distance from the line of sight. Therefore, the two points defining a size are likely to have pixel steps occurring at different distances since they are likely to be different distances from the line of sight.

Given a function, r[x], that rounds a value, x, to the nearest integer and two projected endpoints, P_1 and P_2, we can calculate the projected size, S,

$$S = r[P_1] - r[P_2$$

and the correctly sampled size,

$$S = r[P_1 - P_2]$$

Perceptual Implications:

We have presented four sampling artefacts found in static 3D CGI:

- Inaccurate position
- Inaccurate size
- Inconsistent size
- Inconsistent proportions

The perception of these artefacts is a function of the visual context. For example, if one pixel is large relative to the size of the object, size and proportion inconsistencies are likely to be more noticeable. The ability to perceive a change in size is a function of the ratio of the size of the change to the objects size, a fact later confirmed experimentally. Inaccuracies in the projected size and position of an object limit depth acuity. Changes in depth may not be detectable; objects at different depths may appear identical. Inconsistencies in size also restrict task performance. Obviously, a user cannot compare two objects located at the same distance when one is not visible due to a size rounding error. Similarly, two objects at the same distance may have different sizes or proportions. Clearly, relative depth judgments are severely affected by these artefacts.

According to Johnsons classic study of screen resolution, many simulator tasks require four integral tasks to be performed: target detection, target orientation, target recognition and target identification. All of these tasks are significantly affected by inconsistencies in size and proportions. Target detection is difficult, since the distance presented with the given perspective size and location may differ from the intended distance ([2]). Target orientation is skewed by incorrectly presented proportions, especially for smaller objects or objects at a large distance. Target recognition and identification are similarly affected as these inconsistencies occur.

Multi-polygon objects do not suffer from internal inconsistencies because most graphics routines for complex objects compensate for internal rounding errors. However, they do not correct for the silhouette of the object. This is also seen in textured objects. If two identical, textured objects are being compared, the change in the objects silhouettes causes the original texture to be sampled differently. Even if the one-pixel change in projected size is undetectable, the change in the pattern may be visible.

Textures mapped to two objects differing by a pixel in size.

Therefore, complex objects also suffer from the sampling artefacts described above.

2. Methodology

The importance of a depth cue is a function of the type of task and the viewing parameters. , we are only considering fixed lines of sight that orthogonally intersect the centre of the screen. Therefore, we must choose an appropriate experimental task. Any experimental task should avoid being overly simple while ensuring that noise from any additional complexity is minimized. In the case of immersive CGI, an interactive depth task such as a peg-in-hole, object assembly, or object tracking seems to be suitable ([3]). However, in this chapter, interactive tasks are inappropriate because we want to distinguish static pictorial cues from motion cues.

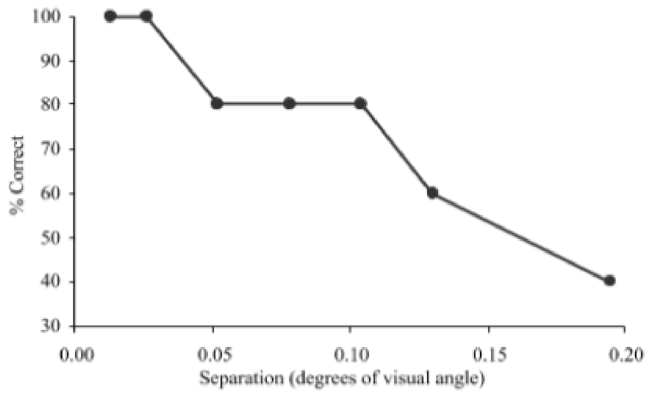
Another experimental task, estimating a single objects depth, requires the addition of familiar size to the depth cues shown to the viewer. Since the artefacts under consideration affect the object's proportions, a familiar object would be distorted and unwanted noise would be introduced into the experiment. Target orientation, recognition and identification tasks also use familiar size cues ([4]). Relative depth estimation (i.e., estimating the distance between two objects on the screen) is prone to large intersubject differences ([5]), although asking subjects to make forced-choice judgments about relative depth reduces these differences.

Holway and Boring first presented the forced-choice methodology to evaluate size-distance relationships [1941]. Since then, many variations on this experiment have been performed ([6]). We also designed a series of relative depth comparison experiments. However, unlike Holway and Borings original work where comparisons were made between objects seen from different viewpoints, the objects to be compared were presented adjacent and simultaneously. The subjects were asked to determine which of two stimuli appeared closer.

Effect of Separation on Location and Size Judgments:

Before we address static perspective depth acuity, we need to understand how the distance between objects affects the detectability of differences in 2D size or location. Spatial acuity decreases as a linear function of the eccentricity from the fixation point (for angles of less than 20°) ([7]). Furthermore, Matsubayashi showed a decrease in depth acuity as a function of separation. Therefore, we need to show how separation affects spatial acuity in a scenario that mimics the one-pixel differences caused by sampled perspective depth information. We performed an informal experiment to provide the preliminary information needed to design future experiments that focused on perspective depth acuity.

In the first set of trials, subjects compared two objects and were asked to identify which was higher on the screen (vertical case) or which was greater distance from the centre of the screen (horizontal case). On the second set, subjects judged which object was wider or taller. The objects differed by a single pixel in all cases. The results show increased separation reduced the detectability of changes in location and vertical size. Horizontal size judgments were not affected by separation.



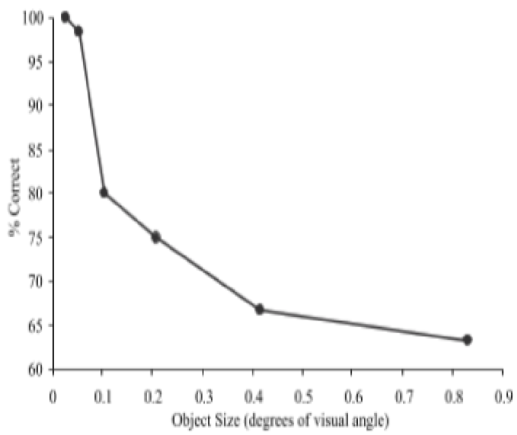
Vertical position acuity as a function of angle separation between the two stimuli. The further apart the stimuli, the more difficult it was to detect a one-pixel difference in the objects vertical location.

This experiment suggests that in typical viewing conditions, an observer can accurately discriminate the vertical position of two objects only when they are less than 0.10° degrees apart ([8]). Therefore, when designing later experiments on perspective depth acuity, we used this threshold to ensure that separation of the stimuli did not affect the relative judgment of depth.

Effect of Size on Location and Size Judgments:

The other factor we expect to influence the ability to distinguish one-pixel differences in projected size and location was the size of the object. For a large object, a one-pixel change in size or position may only represent a small change relative to the size of the object, while for a small object, a one pixel change may be more significant. We conducted an experiment to see if this ratio influenced the ability to detect one-pixel changes in projected size and position. We used the same experimental

task as the previous experiment. The results showed that decreasing the objects size increased the accuracy of judgements about one-pixel differences in both size and location.



Pilot experiment data showing the effect of the vertical and horizontal size of the object on the detectability of a one-pixel change in vertical size. The smaller the object, the more easily subjects distinguished a one-pixel difference in vertical size.

This experiment suggests that a viewer can typically discriminate one-pixel changes in vertical location and size for objects subtending less than 0.40° of visual angle. Using the ratio of pixel size to object size as a different metric, the threshold at which one-pixel changes in vertical size and location were detectable occurred when the pixel size was $1/16^{th}$ of the object size. Similar results were found for changes in horizontal size, although changes in horizontal position were difficult to detect even for very small objects. We designed the stimuli in later experiments to be small enough that changes in projected size and location would be detectable.

Detectability of Sampled Perspective Cues:

When changes in position and size in both dimensions represent different perspective depths, we call them *sub-cues* of linear perspective ([9]). Having established how 2D size and separation affect the detection of changes in size and position, we can now investigate the detectability of changes in depth due to one-pixel changes in the four perspective sub-cues. One-pixel steps in the sub-cues do not occur independently. All combinations of sub-cues are possible if we vary the 3D position and size. This leaves us with 2^4 possibilities, from no change in location or size, to a one-pixel difference in all four sub-cues.

A formal experiment was conducted to determine how the ability to detect differences in depth varies with the 15 different possibilities in which a pixel step occurs. Pixel size was also varied. Subjects judged which of two objects in a scene appeared closer as different possible combinations of sub-cues were presented. The size and separation of the objects were chosen based on the results of the two previous informal experiments. Details of the experiment design and analysis of the results are presented in Experiment A. The main results can be summarized:

- Decreasing pixel size reduced detectability of a fixed change in depth
- Vertical position sub-cues significantly increased accuracy when presented individually; the other sub-cues did not
- Combining sub-cues increased the detectability of changes in depth

For two points of interest, an optimal viewpoint can be found by changing the viewpoint such that the vector describing the difference between the two points maps to one of the diagonals of the screen. The *axis-angle method* manipulates the viewpoint as follows:

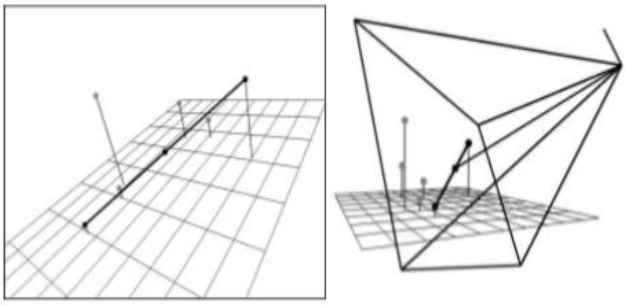
a. Translate the viewer so that the first point, P_1, is at the origin
b. Calculate the axis of rotation, R, and angle, β, for $\vec{E}$ such that $\overrightarrow{P_1P_2}$ lies on the diagonal of the screen, $\vec{V}$:

$$\vec{R} = \vec{V} \times \overrightarrow{P_1P_2}$$

$$\beta = \cos^{-1}(n[\vec{V}] \cdot n[\overrightarrow{P_1P_2}])$$

c. Rotate the viewpoint around the axis, $\vec{R}$, by the angle, β
d. Translate $\vec{E}$ by the viewing distance, *d*, and centre the vector in the screen, *c*.
e. Scale the viewpoint by *s* to match the diagonal of the screen

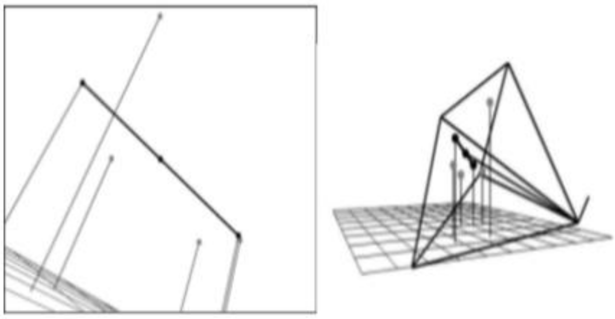
Viewpoint manipulation method for maximizing the distance between two points.



An example of manipulating the viewpoint to maximize distance between two points. The left image shows the projected image. The right image shows the location and orientation of the viewing frustum.

Viewpoint manipulation has some disadvantages. Given an entire scene, manipulating the viewpoint

with the axis-angle method could adversely effect other important information, such as the viewers sense of self-location.



Viewpoint manipulation can result in lost information about the rest of the scene.

Some heuristics can be applied to maintain certain characteristics of the viewpoint. If we wish to maintain a sense of up, the rotation can be limited to yaw without any disorientating roll or pitch. The vector can then be scaled to match the diagonal of the screen. The *yaw-scale method* manipulates the viewpoint as follows:

With the axis-angle method could adversely affect other important information, such as the viewers

a. Orbit the line of sight, $\vec{E}$, around the y-axis by the angle, β, calculated using the x and z components of $\overrightarrow{P_1P_2}$:
b. Scale the scene by (s_x, s_y) so that the projection of $\overrightarrow{P_1P_2}$ lies maximally on the diagonal

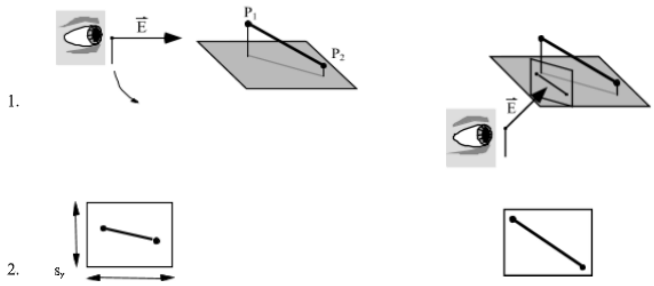
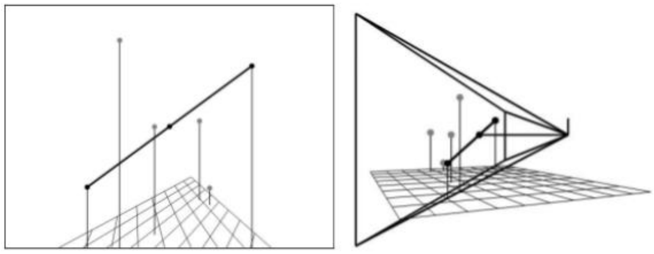
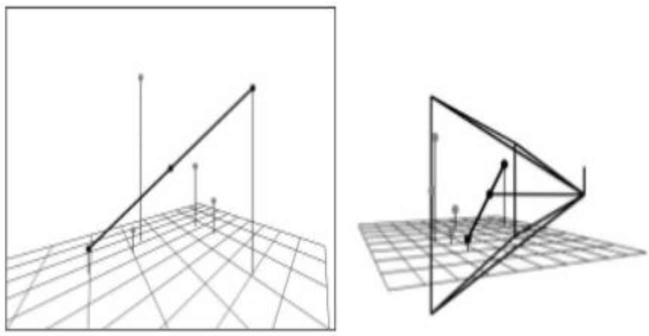


Figure 5.18: Manipulating the viewpoint using only rotation around the y-axis and scaling to fit the vector to the diagonal.

Restricting the range of movement of the viewpoint with the yaw-scale method does not affect the result; the difference between the two points is again maximally presented.



The yaw and scale viewpoint manipulation method.



The disadvantage to this method is that the scaling is not aspect-constrained; it can warp other objects in the scene.

The yaw-scale method can distort the scene.

However, using a bounding volume to ensure all objects of interest are present on the screen can greatly reduce this distortion. Similarly, maintaining the aspect ratio when scaling can also remove unwanted distortions.

3. Conclusion

This chapter has described and evaluated the effects of sampling on static linear perspective depth cues in 3D CGI. Rounding errors in the projected size and position of a 3D object cause inaccurate representation of depth. These errors also result in inconsistently presented size and shape. Experimentation demonstrated that these artefacts influence relative depth judgments in some situations. Small objects, close to the line of sight and a large distance from the viewer are the most susceptible.

Two methods for alleviating these artefacts have been presented. Manipulating the projected position of objects endpoints ensures an objects size is consistently presented. Moving the viewpoint to an optimal location maximizes the number of pixels differentiating points of interest. Both of these methods are computationally inexpensive for static scenes.

References

[1] Kalawsky, R. 1993. *The Science of Virtual Reality and Virtual Environments*. New York: Addison Wesley Publishing.
[2] Pausch, R., Proffitt, D. & Williams, G. 1997. "Quantifying Immersion in Virtual Reality. In *Proc. SIGGRAPH '97*, pp. 13-18.
[3] Pfautz, J. 1996. *Distortion of Depth Perception in a Virtual Environment Application*. Masters Dissertation, Massachusetts Institute of Technology. Pfautz, J. & Robinson, P. 1999. An Analysis of Perspective Geometry and Antialiasing. in*Proc. Eurographics UK 17th Annual Conference*.
[4] Smets, G. & Overbeeke, K. 1995. Visual Resolution and Spatial Performance: The Trade-Off between Resolution and Interactivity. in *Proc. VRAIS '95*, pp. 67-73.
[5] Johnson, J. 1958. Analysis of Image Forming Systems. In *Proc. Image Intensifier Symposium*, pp. 249-273.
[6] Barfield, W. & Rosenberg, C. 1995. Judgments of Azimuth and Elevation as a Function of Monoscopic and Binocular Depth Cues Using a Perspective Display. *Human Factors*, 37 (1), pp. 173-181.
[7] Baird, J. 1970. *Psychophysical Analysis of Visual Space*. London: Pergamon Press..
[8] Graham, C. 1951. Visual Perception. in S. Stevens (Ed.), *Handbook of Experimental Psychology*, pp. 868-920. New York: John Wiley & Sons.
[9] Pfautz, J. 1996. Distortion of Depth Perception in a Virtual Environment Application. Masters Dissertation, Massachusetts Institute of Technology. Pfautz, J. & Robinson, P. 1999. An Analysis of Perspective Geometry and Antialiasing. In Proc. Eurographics UK 17th Annual Conference.