VIEW DEPENDENT PERSPECTIVE IMAGES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Jeremiah Fairbank

August 2005

ABSTRACT

The effects of scalar distortions that result from incorrect viewing location adversely affect a viewer's ability to judge space and scale in a perspective image. A study of the history of the development of the perspective image shows that artists and architects of the Renaissance were sensitive to the location of the viewer. Using this precedent, we place a renewed emphasis on the relationship between the viewer of an image, the image, and the scene the image represents.

We have created a computer graphics application that uses a skew camera model to present a view-dependent perspective image. We find the viewer's location using a camera based tracking system. Viewer location data, along with information about display location guide the skew camera's parameters. We create perspective images that do not suffer from the distortions that are commonly present in wide-angle perspective images. Our system allows for the display of perspective images that geometrically recreate what the observer would see if they were to view a real scene.

Biographical Sketch

The author was born in Ithaca, New York on July 7,1978. After completing architecture school at Cornell he continued study with the Program of Computer Graphics.

Acknowledgements

There are many without whose support this project would not have been realized. Don Greenberg was the impetus for all of this. The effect of his academic, professional, and personal guidance along with his zeal cannot be described. We do indeed ”stand on the shoulders of giants.” My family has always supported me in whatever I do, so this endeavor should not have been expected to be different in that regard. It was not. The members of the program of computer graphics all helped me to complete this work, I name a few in particular: Alex Lieberman provided help with math and coding. We worked together on another project, the code from which I use here to import meshes. I shared an office the first year with Jacky Bibliowitz and Adam Kravetz. They helped me to get up to speed on this newer side of campus. Vikash Goel helped with camera tracking. Hurf Sheldon makes things work.

Table of Contents

1 Introduction 1

2 Perspective Illustrations in the Renaissance and Beyond 3
- 2.1 Why Look Back? 3
- 2.2 An Early History of Perspective 4
- 2.3 Development of the Perspective Image 5
  - 2.3.1 Perspective Machines 11
  - 2.3.2 Masaccio Trinity 14
- 2.4 Renaissance Sensitivity to Viewer Location 17
  - 2.4.1 Three-Dimensional Renditions that are Sensitive to Viewer Location 20
  - 2.4.2 Surface Renditions that are Sensitive to Viewer Location . . 27
- 2.5 Marginal Distortion 30
  - 2.5.1 Curvilinear Perspective 40
- 2.6 History Conclusion 43

3 The Perspective Image 47
- 3.1 Picture Perception 47

4 View Dependent Perspective Images 57
- 4.1 Introduction to the Problem 58
- 4.2 User Location Tracking 58
- 4.3 Camera Based Tracking 60
- 4.4 Skew Camera 63
- 4.5 Tiling Displays and Networking 72

5 Conclusion 75

A Depth Perception 77
- A.1 Interposition 77
- A.2 Relative Size 78
- A.3 Height in the Visual Field 78
- A.4 Motion Parallax 78
- A.5 Binocular Disparity 78

A.6 Accommodation and Convergence 79
A.7 Size Constancy 79
A.8 Aerial Perspective 79

B User Location Tracking 80
B.1 Mechanical Tracking 80
B.2 Inertial Tracking 81
B.3 Magnetic Tracking 82
B.4 Acoustic Tracking 82
B.5 Optical Tracking 83
B.6 Ultra Wide-Band Tracking 83

C Digital Perspective Images for Architectural Design 85
C.1 Design 85
C.2 The Perspective Transform 90

Bibliography 93

List of Figures

2.1 Stroebel: Portrait 7
2.2 Magritte: The Human Condition 8
2.3 Brunelleschi: The Florentine Baptistry 9
2.4 Baptistry experiment recreation 10
2.5 daVinci: sketch 11
2.6 Dürer: perspective machine 1 13
2.7 Dürer: perspective machine 2 14
2.8 Cigoli: perspective machine 15
2.9 Masaccio: Trinity 16
2.10 Masaccio: Trinity diagram 18
2.11 Masaccio: Trinity analysis 19
2.12 Borromini: Palazzo Spada plans 21
2.13 Borromini: Palazzo Spada 21
2.14 Borromini: Palazzo Spada 22
2.15 Bernini: Scala Regia 23
2.16 Bramante: Church of Santa Maria San Satiro plans 24
2.17 Bramante: Church of Santa Maria San Satiro 25
2.18 Teatro Olympico 25
2.19 Teatro Olympico 26
2.20 Hans Holbein: The Ambassadors 28
2.21 Apartment of the Gesu cherubs 28
2.22 Apartment of the Gesu ceiling 29
2.23 Pozzo: San Ignazio dome 31
2.24 Pozzo: San Ignazio vault 31
2.25 Pozzo: San Ignazio diagram 32
2.26 Peruzzi: Villa Farnesina 33
2.27 Uccelo: memorial to Sir John Hawkwood 36
2.28 Off-axis sphere diagram 37
2.29 School of Athens 38
2.30 Michelangelo: David 39
2.31 The problem of wide-angle linear perspective 41
2.32 William Herdman: Curvilinear Perspective 42
2.33 Guido Hauk: the conform system 44
2.34 Schoellkopf panorama 45

3.1 Depth perception graph 48
3.2 Kitchen overlay 50
3.3 Camera comparison diagram 1 52
3.4 Camera comparison diagram 2 53
3.5 Skew camera comparison 54
3.6 Image stream 55
3.7 Image stream comparison 56

4.1 Camera tracking 61
4.2 Infrared camera 62
4.3 Tracking camera screenshots 63
4.4 Skew camera 65
4.5 Skew camera comparison 66
4.6 Standard camera 67
4.7 Standard perspective transform 71
4.8 Skew camera perspective transform 71
4.9 Network 72
4.10 Display photo 73
4.11 Display comparison 74

C.1 Rand Hall freshman design studio 88
C.2 Rand Hall student computer lab 89
C.3 Architectural perspective construction 91
C.4 Probe 92

Chapter 1

Introduction

Our perception of the space we inhabit is based on the light that enters our eyes. This unique collection of light rays that enter the eye to form a retinal image has been described as a "bundle of rays." [FR80] Perspective images, which can geometrically recreate what a viewer might see if they were looking at a real scene, are powerful tools for visualizing three-dimensional space. But perspective images are only two-dimensional. If a perspective image is to correctly recreate our perception of the real scene, the "bundle of rays" that it sends to an observer must replicate those that would reach the observer from the real scene. This implies knowledge of the observer's viewpoint. When viewed from an incorrect location, the image may appear distorted. This is because the "bundle of rays" that the image sends to the observer's eye does not exactly correspond to those that would reach the eye from a real scene.

The effects of scalar distortions that result from incorrect viewing location adversely affect a viewer's ability to judge space and scale in an image. If our goal is to provide a better sense of space and scale when viewing perspective images, we must account for these distortions.

A study of the history of the development of the perspective image guides our implementation of computer graphics applications. Historically, perspective images were often used to extend space by giving an illusion of greater size or depth. The illusions required that the observer be in the correct viewing location. The artists and architects of the Renaissance were sensitive to the location of the viewer, certainly more sensitive than the projection techniques we use today. We place a renewed emphasis on the relationship between the viewer of an image, the image, and the scene the image represents.

We have created a computer graphics application that presents a view-dependent perspective image. We find the viewer's location using a camera based tracking system. Viewer location data, along with information about display location guide the skew camera's parameters. Our application uses a skew camera model to draw a unique perspective image that is based on the location of a viewer. Each skew camera takes into account the location size and orientation of the display. Multiple displays can be tiled to create a wide-angle perspective image that does not suffer from distortions that are typically inherent in the display of wide-angle images.

Our goal is to create perspective images that do not suffer from the distortions that are commonly present in wide-angle perspective images. We do this by using a skew camera that uses viewer location data to present an image that geometrically recreates what the observer would see if they viewed a real scene.

Chapter 2

Perspective Illustrations in the Renaissance and Beyond

This chapter describes the historical precedent for the creation of view dependent perspective images. We discuss examples in which there was a known spatial position of images (or scenes) and observers. These images were constructed so that they gave the impression of being a ”window into a three-dimensional world.” Frequently the images were used to extend space by giving an illusion of greater size or depth.

2.1 Why Look Back?

We attempt to solve a modern-day problem in computer graphics: the relationship between the viewer or designer and the computer generated image. There exists a wealth of information in historical record that details how others have dealt with similar issues. The history begins long before the invention of the computer as we know it. A logical place to begin is with the early invention and formulation of perspective. In doing this, we will show that historically others have been more sensitive to the role and location of the viewer of art or architecture. We will use this study of history to guide the implementation of new computer graphics applications that will better provide the experience of viewing a virtual three-dimensional world, and thus better suit the needs of designers.

2.2 An Early History of Perspective

Much of the development of perspective images as a means for representation occurred during the Italian Renaissance from the 15th to the 17th centuries, but its formulation began even centuries earlier as philosophers began to understand the foundations of human perception. As such, perspective can trace its history to as early as the fourth century B.C. when Euclid[1] used the terms *visual ray* and *visual cone* to describe how light from a scene enters the eye. An understanding of visual perception later developed, but was based on Euclid's theory that rays of light entering the eye determine what people see. [2]

Before the Renaissance, the use of perspective as a representational tool was tangential to the the primary focus of study for a variety of philosophers and scientists. Even though they were not personally interested in using what they had learned for representational purposes, their discoveries formed the base upon which

---

[1] optika

[2] Plato and Aristotle argued about the source of Euclid's rays of light, but eventually came to the same conclusions. In *Timaeus*, Plato claims that a *firelight* originating inside the eye enters a scene, travels forward until it strikes an object, then returns to the eye with information about surface lighting. While fundamentally flawed in his understanding of the origin of light rays, Plato's description is interesting now as a description of raytracing in modern computer graphics. His student Aristotle reversed this idea and correctly claimed that the eyes do not produce light, but receive light from the surrounding environment.

the Renaissance could flourish. Euclid and Ptolomy focused on the study of optics. Plato and Aristotle interested themselves primarily in philosophical theories of light. Pythagoras examined light because of an interest in geometry. Galen contributed to the knowledge of perception as it related to physiology and anatomy[3]. Several hundred years before the Renaissance, Ibn-al-Haitham collected, deepened and combined many of these theories[4]. He claimed that light rays came to the eye in the shape of a pyramid, and that the eye's surface acted as a receptor to any light impulses emitted from any point in the pyramid. All these inquiries focused only secondarily, if at all, on the creation of perspective images. It was not until the Renaissance that an intersection occurred between developments in optical theory and art and architecture[Bor04]. Based on the process of human perception, as it had been described for centuries before, Renaissance artists and architects gave the viewers of pictures a greater sense that they were viewing a real scene. This example of interdisciplinary innovation led to the use of perspective images as a method of representation.

2.3 Development of the Perspective Image

Early philosophers knew that the sensation of vision came about because rays of light entered the eye from a scene. Renaissance inventors of perspective theorized that if an image were to send the same set of rays to the eye as would a real scene, the viewer would get the impression that they were viewing a real scene. This concept is demonstrated in a modern picture by Leslie Stroebel in which the photographic image seems to blend into the surrounding scene [SCCZ86] (Figure

---

[3] *De Usu Partium*

[4] Ibn-al-Haitham is also often referred to as Alhazen

2.1). Stroebel's demonstration is but the latest in a long series of similar examples of the same phenomenon. For example, a 1961 psychology experiment showed that observers could not tell the difference between a real scene and an image of the scene as long as the image and scene were viewed from the same monocular viewpoint [Kub86]. A similar approach was taken by the surreal artist Rene Magritte. In his painting *The Human Condition*, Magritte shows how a painting might act as a window into another world [Tor77] (Figure 2.2).

The examples above are only modern re-enactments of the method used by Brunelleschi in the early 15th century to depict the Florentine Baptistry. In his experiment, Brunelleschi set up his easel inside the main door of the Florentine Cathedral in order to make an image of the adjacent baptistry (Figure 2.3). A peephole in the painting and an appropriately placed mirror allowed him to fix the viewing location and to restrict the observer to monocular viewing (Figure 2.4).[5] For these reasons, the image Brunelleschi created fit within the context of the real scene. The ability to create an image on a two-dimensional surface that gives the viewer the impression that they are viewing a real scene had great potential during the Renaissance, not only as a novelty, but as a representational tool for artists and architects. Frequently the aim of these artists was to use their art as an extension of perceived space.[6] The invention of the perspective image was based on centuries of knowledge about human perception, but it was not until the Renaissance that it saw widespread application to art and architecture.

---

[5] For a more detailed technical description of the method Brunelleschi used to create the Florentine Baptistry image, see appendix II of [Kem90]. The original source description of this experiment can be found in [Man70].

[6] this phenomenon is also evident in wall paintings discovered at the ruins of pompeii that depict an extension of space with a perspective painting of columns.

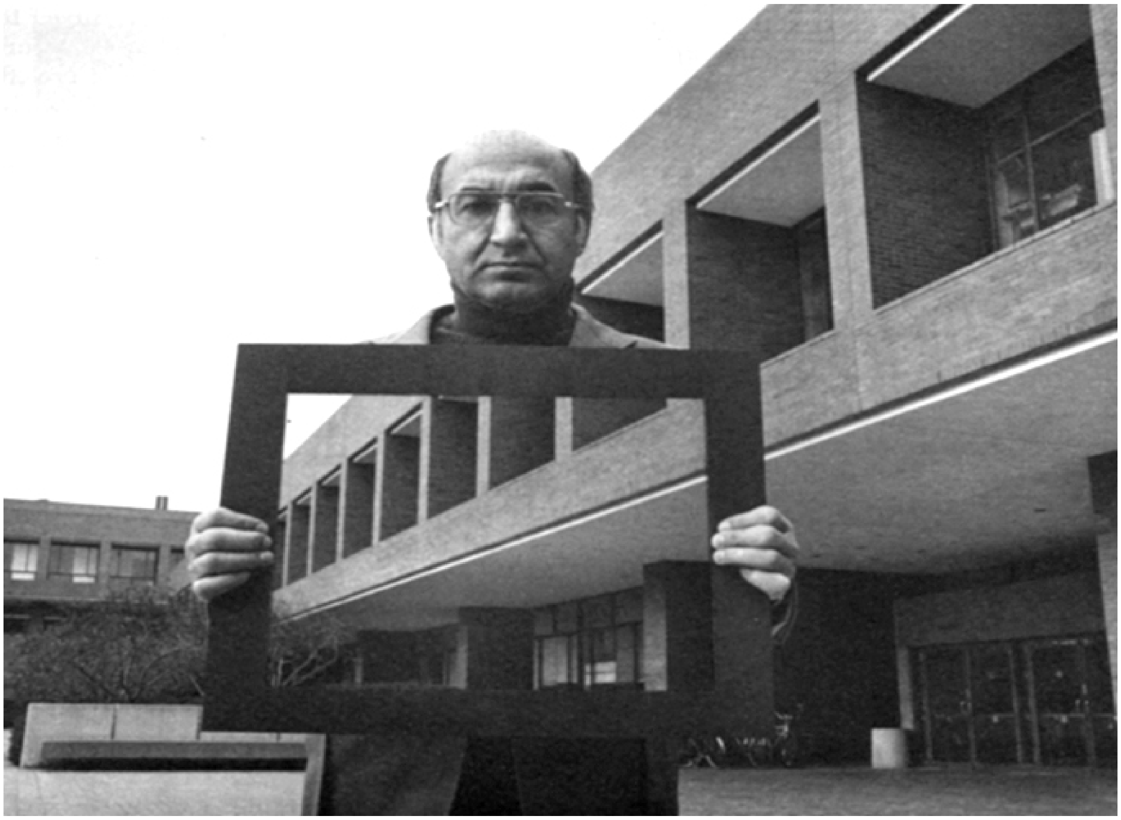


Figure 2.1: A portrait by Leslie Stroebel demonstrates that a perspective image, in this case a photograph, can give a viewer the impression that they are viewing a real scene. By precisely locating the camera and print with relation to the surrounding scene, the light entering the eye from the print the subject is holding matches that which would enter the eye from the occluded scene.

Figure 2.2: Rene Magritte's painting ***The Human Condition*** demonstrates how a painting could act as a window into another world. Because the painting reflects the same "bundle of rays" toward the eye that the occluded scene would, it fits into its surroundings as if it were part of them.

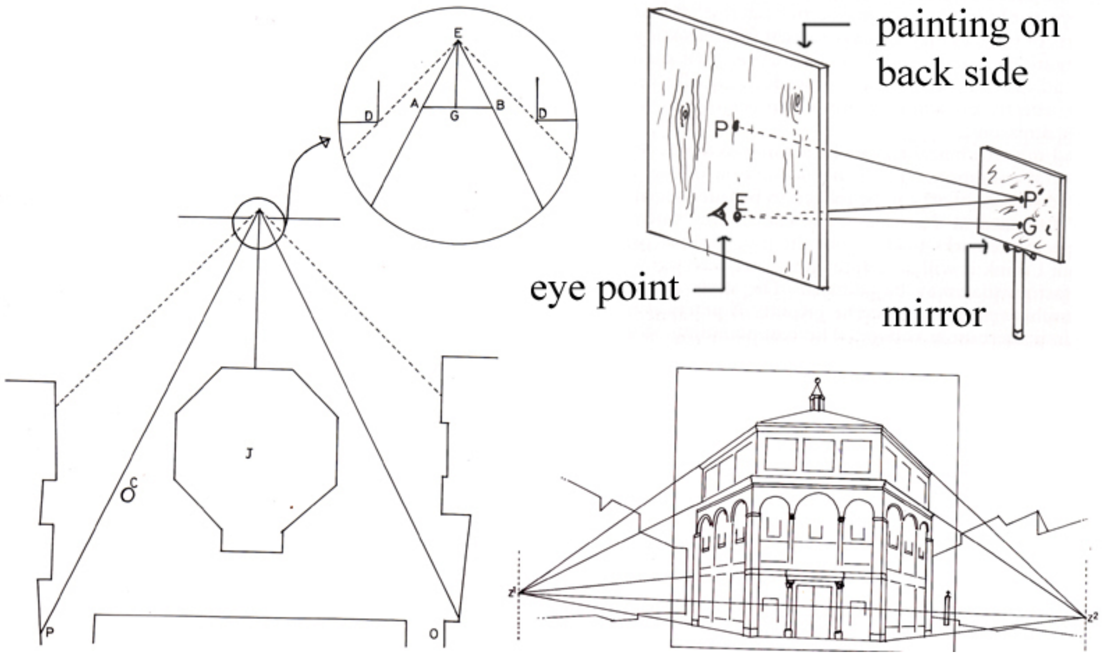


Figure 2.3: ”Brunelleschi’s peepshow” was an experiment conducted in the early 15^{th}century and was a pioneering example of applying what was known about human perception to the creation of perspective images for the representation of a scene. At left is a plan of the Baptistry and the location at the cathedral entrance from which the image was created. At the top right is a drawing of the painting/mirror setup. At the bottom right is the final image inset on a wider view of the plaza. Notice that the perspective in the image matches that of the scene surrounding it.

Figure 2.4: A recreation of Brunelleschi's experiment at the entrance to the Duomo in Florence facing the piazza and baptistry. A perspective image (reflected in a mirror) of the baptistry is seen through a small hole in the board. Correct placement of the image and mirror make it possible for the perspective image to exactly replace what the viewer would see if the mirror were removed.

2.3.1 Perspective Machines

The development of architectural design and representation during the Renaissance was closely linked to the development of perspective as a tool for seeing space. During the Renaissance, many artists and designers understood the space around them as if seen through a picture plane. Many early demonstrations of perspective present the picture plane as a literal window through which one sees the world. Leon Battista Alberti claimed that one way of aiding the architect in showing three-dimensional space on a two dimensional surface is to observe the scene through a thin viel.[7] Similarly, early sketches by Leonardo daVinci show the use of a transparent plane to draw a sphere (Figure 2.5). Leonardo also warned against

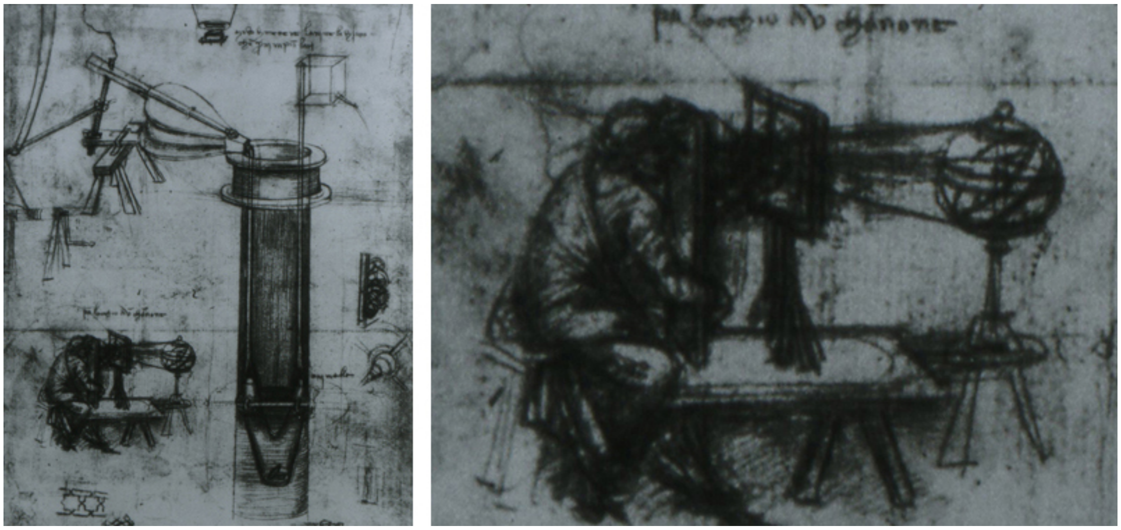


Figure 2.5: A viewer observes a sphere through a transparent plane. The image of the sphere drawn on the plane is a representation of the sphere as it is seen from a fixed station point. In order to fix the eye, the viewer looks through a hole in a vertical board. Leonardo daVinci *Draughtsman Using a Transparent Plane to draw a sphere, 1510.*

---

[7] In his text *della Pittura (1453)* Alberti refers to a thin *velo*, a semi-transparent cloth upon which one could simultaneously paint a scene while observing the same scene through the *image plane.*

the use of such a setup to the exclusion of true skills in design and representation:

> "There are some who look at things produced by nature through a glass, or other surfaces or transparent veils. They trace outlines on the surface of the transparent medium… But such an invention is to be condemned in those who do not know how to portray things without it, nor how to reason about nature with their minds… They are always poor and mean in every invention and in the composition of narratives, which is the final aim of this science." [Kem90]

This warning, along with the accompanying diagrammatic sketch, bears striking relevance for those who would now use perspective images, and particulary computer graphics images, as a tool for seeing and understanding space. Beyond the obvious stated need for designers to cultivate their perceptual and representational skills lies a fundamental concept: In creating perspective images, there is a specific relationship between the viewer, the image, and the scene the image represents.

Renaissance inventions such as those by Albrecht Dürer (Figures 2.6 and 2.7) and Ludovico Cigoli (Figure 2.8) illustrate the roots of such a concept; that perspective images of a scene should be viewed from a given station point. In both cases, points from a three-dimensional scene trace back to a fixed eye or station point. At the location where the line intersects an image plane, the operator registers the point's two-dimensional representation. Lines then connect points to create perspective images of shapes and eventually entire scenes. These machines demonstrate that the principles used to invent perspective lent themselves to the creation of mechanical aides. When used properly, the aides helped designers to portray in two dimensions what actually existed in three dimensions.

Figure 2.6: A woodcut by Albrecht Dürer shows the use of a hinged plane upon which the user registers the points of a three-dimensional scene. Each point is traced from the scene at left to the eye at right. The artist registers a point where the string intersects the hinged image plane.

Figure 2.7: A second machine for the construction of perspective images by Albrecht Dürer. This version replaces the hinged opaque plane with one that is fixed and transparent. While the artist sits close to the transparent plane, the *eye* position or station point is where the string is fixed to the wall.

2.3.2 Masaccio Trinity

The previous examples all demonstrated the use of perspective to mechanically record an existing scene as a perspective image. Masaccio's fresco of the Trinity is an early example of the use of perspective to represent in two dimensions a *non-existent* three-dimensional scene (Figure 2.9). This was significant because designers began then to more frequently create images of unbuilt scenes. This advance opened the door for the use of perspective in art and architecture for showing *what might be* as opposed to merely recording *what was*.

Because a fresco is painted in plaster, its location is fixed, as is the ideal location for its viewing. The Trinity fresco's coffered ceiling and colonnaded room are based on and represent some non-existent space. While an infinite number of possible geometries exist that could have been used to create the same image, one is thought to be ideal. We do not have direct access to the original geometric source, but that source can be derived, with certain reasonable assumptions, by

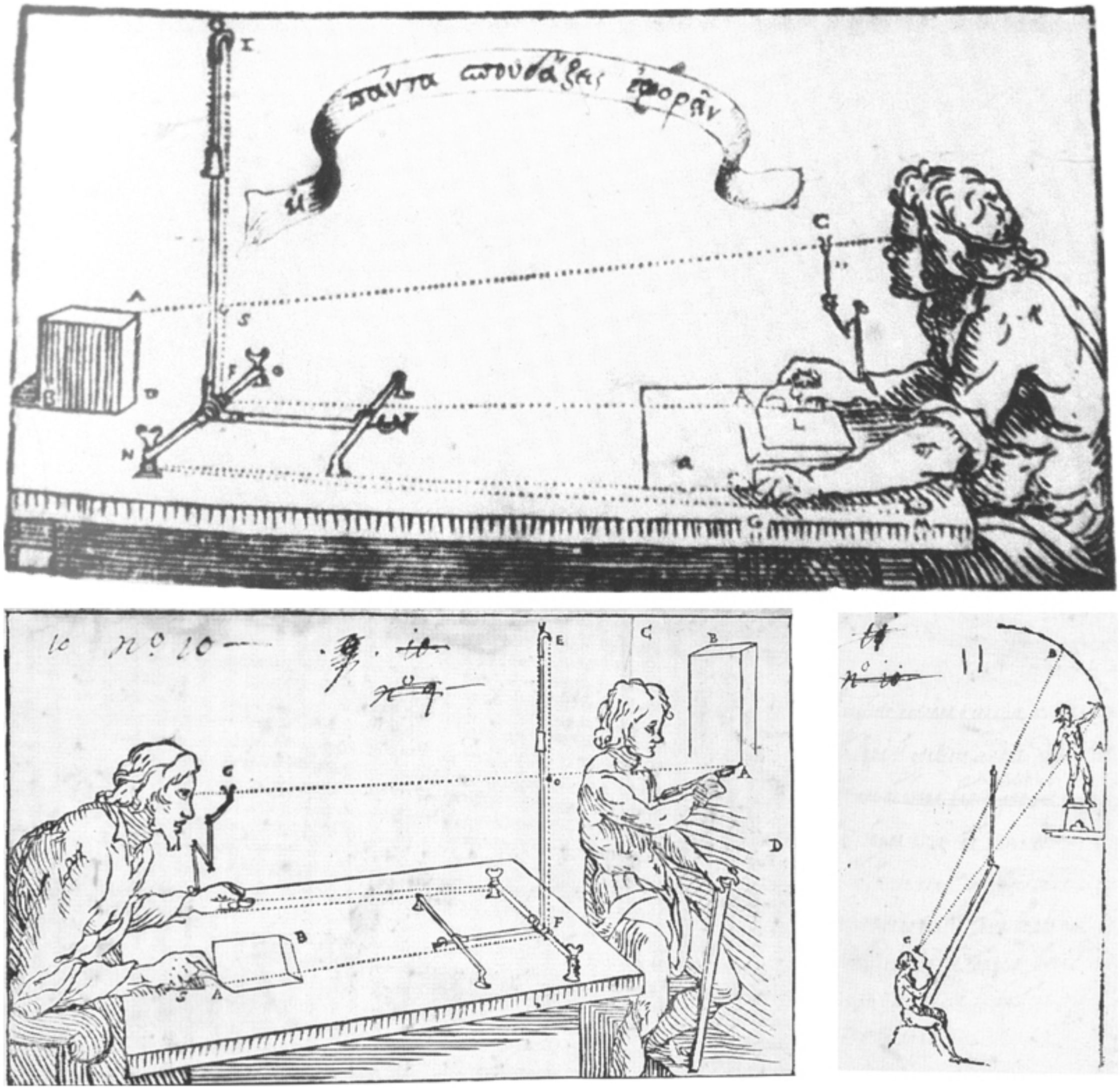


Figure 2.8: A machine from Ludovico Cigoli's *Prospettiva Practica* that can be used for the recording and projection of perspective images. Such a machine could be used for the display of perspective images on arbitrary planes and curved surfaces.

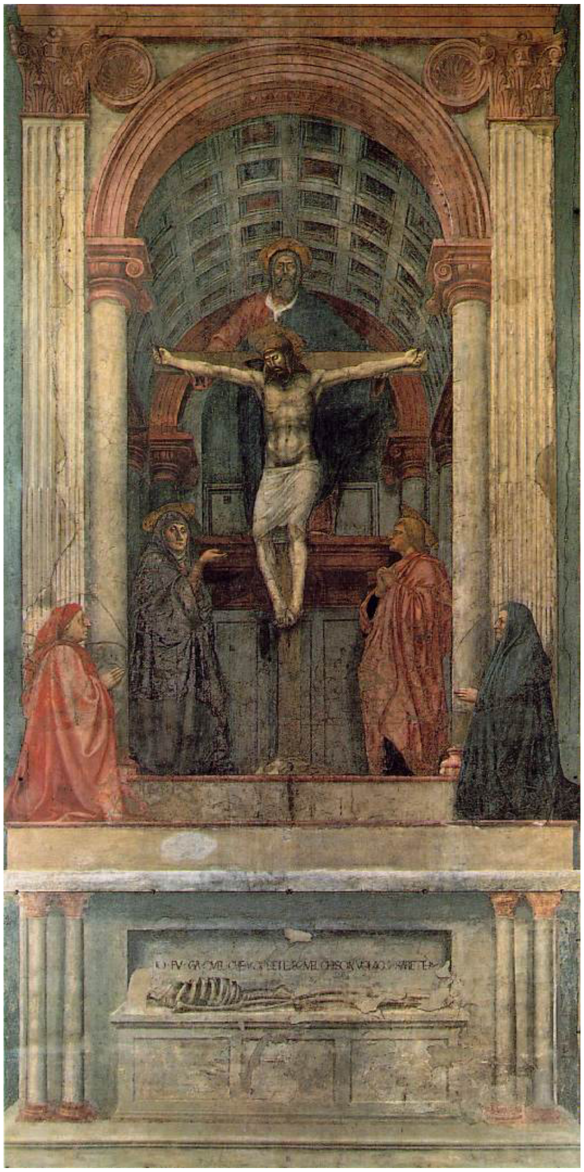


Figure 2.9: Masaccio's Trinity fresco is an early perspective image created based on a non-existent three-dimensional space. By definition, the fresco is always located in the same place on the wall. This makes it necessary that the viewer assume the correct station point, standing at a certain location in front of the fresco, in order to be in the correct viewing location.

reversing the process that might have been used to create the image in the first place (Figure 2.10). A line is traced from the station point through the fresco and out into infinity. The location of the point in the scene geometry that corresponds to the point in the fresco must lie along this line. This process is repeated for each point in the fresco. In this case, the *ideal geometry* is chosen from the infinite number of possibilities based on the preservation of right angles in the source geometry. It should be emphasized that this reconstruction is valid only based on the assumption that the viewer's location is fixed relative to the location of the fresco. In reality, while the fresco is fixed to the wall, the viewer is not necessarily fixed to the ideal station point. After the viewer moves from the ideal station point, a static fresco image represents geometry that is possible but implausible because no ideal regular geometry exists (Figure 2.11 *left*). For the ideal geometry, if the perspective image were to dynamically respond to the viewer's change in location, it would present the correct image based on the new station point (Figure 2.11 *right*). This type of speculation is purely academic. No such ability to create *dynamic frescos* existed during the Renaissance. However, the illustration demonstrates the relationship between the station point used to construct a perspective image, and the location of the viewer of the resultant image. Renaissance designers understood this relationship and made it an integral part of their art.

2.4 Renaissance Sensitivity to Viewer Location

Because Renaissance artists and designers were responsible for the re-discovery of perspective, they also possessed a particular sensitivity when it came to the display of art and architecture. They paid particular attention to the presentation of art or architecture based on the assumption of a known station point. These designers

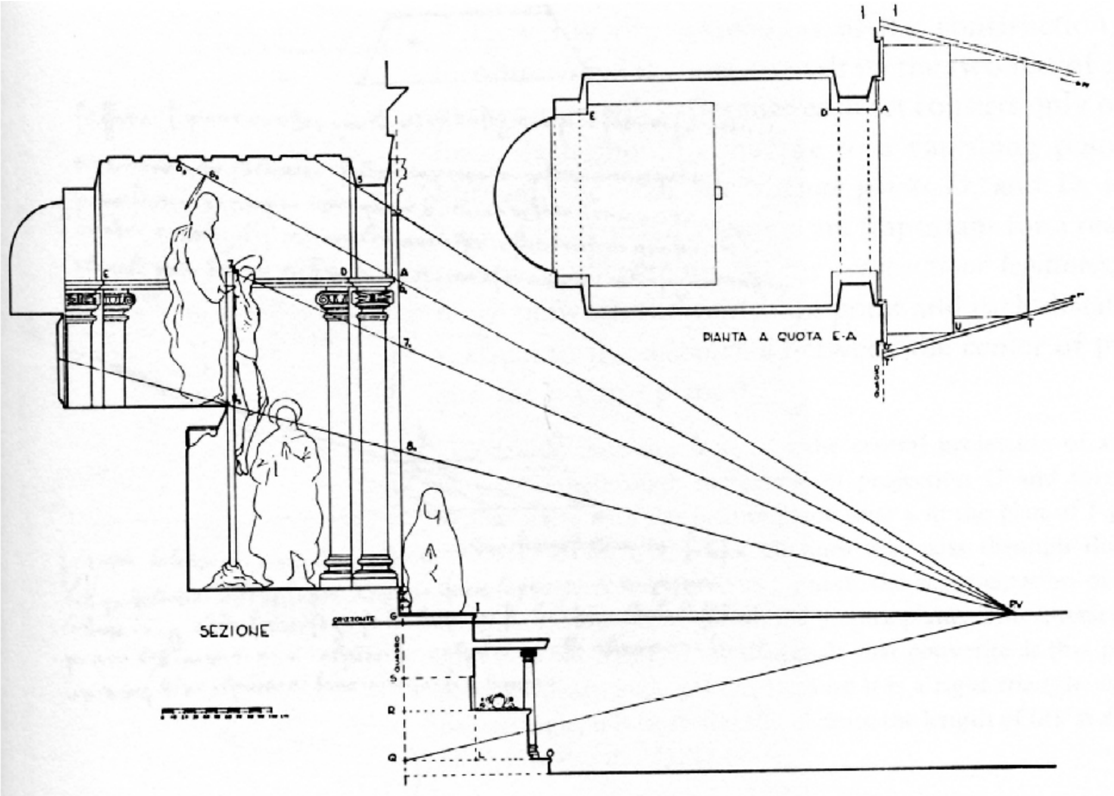


Figure 2.10: An analysis of the Trinity fresco shows the ideal regular geometry that the fresco represents from a given station point. This analysis begins with the image and traces into the scene to reconstruct what may have been the room on which the image was based.

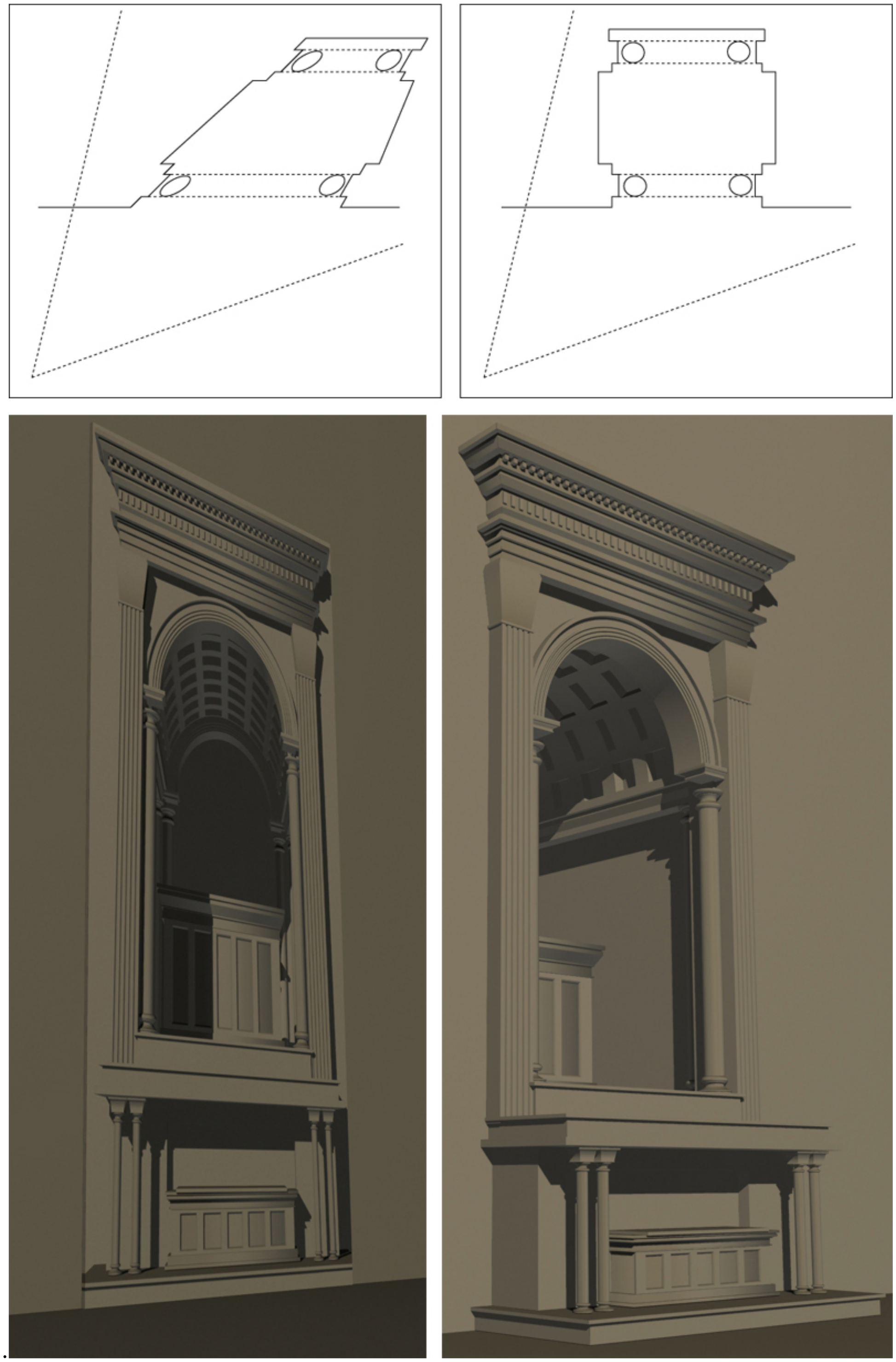


Figure 2.11: Two off-axis analysis images of the Trinity fresco by Masaccio. The image on the right is based on an ideal regular geometry with right angles. the image on the left has no possible ideal regular geometry.

knew the position from which their work would be viewed. They often used this to their advantage by creating an image that would ideally represent a scene for a viewer from a given location. Their goal was to extend the space a viewer perceived beyond what actually existed.

2.4.1 Three-Dimensional Renditions that are Sensitive to Viewer Location

An example of architecture that is sensitive to viewer location is the Borromini Gallery at the Palazzo Spada in Rome. In the 17^{th} century, Cardinal Spada purchased the palazzo and asked Borromini to renovate it by adding an extra colonnade and courtyard garden. Unfortunately, there was not enough space for the grand colonnade Borromini envisioned (Figure 2.12). In order to overcome this obstacle, the architect carefully warped the geometry of the built gallery in order to give the impression of a long colonnade(Figure 2.13).[Tre99] If one views the colonnade from the glass wall of the adjacent library, the impression is of a long colonnade leading to a life-size statue in an expansive courtyard garden (Figure 2.14 *left*). When viewed from any other location, one realizes that the gallery is much shorter and the statue much smaller than was previously expected. This effect was possible because the designer carefully altered the scene according the the rules of perspective. The result was a geometry that, when viewed from the wrong location, appeared warped (Figure 2.14 *right*).

Bernini used the same geometry warping technique that Borromini did, although with more subtlety, in his design for the Vatican's Scala Reggia (Figure 2.15). After applying the required transformation to the geometry, the colonnade and stair did not require the same level of compression that space limits imposed

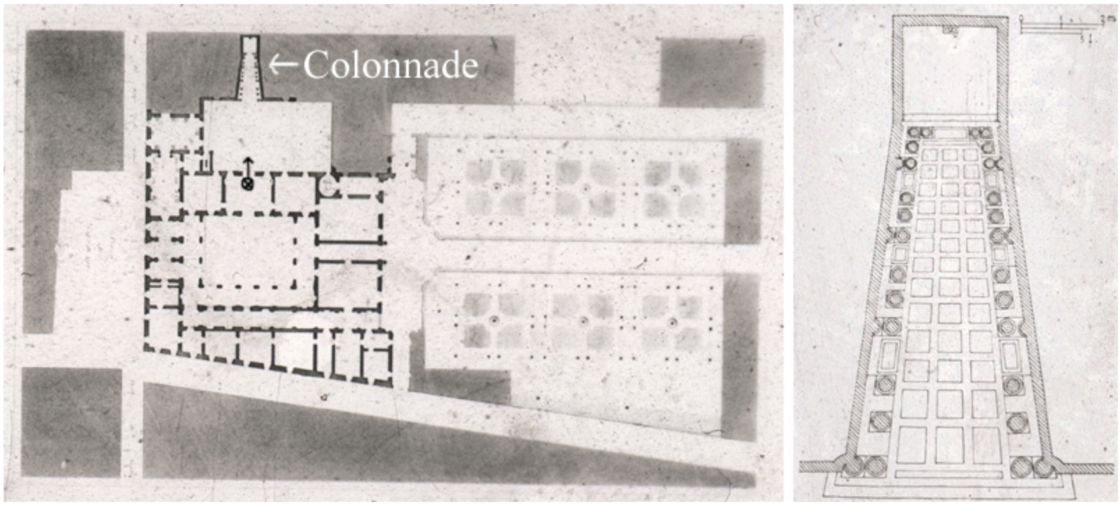


Figure 2.12: Plans of the Borromini Gallery at the Palazzo Spada in Rome show the warped geometry of the colonnade. It is built so that from the correct vantage point it seems much deeper than it really is.

Figure 2.13: These images of the Borromini gallery show the effect of increased depth that was caused by the warping of the scene geometry. At left is an image in which the gallery seems like an extensive colonnade. At right, with the addition of a familiar scale, the illusion is made evident.

Figure 2.14: Images of the Borromini gallery show the colonnade as seen from the ideal station point *(left)* and from an off-axis view *(right)*. Note the warped geometry which creates the illusion of more depth than actually exists in the scene.

on the Borromini Gallery. Note that in both cases the geometry could have been substantially more compressed. The extreme example of this compression would result in a two-dimensional perspective image.

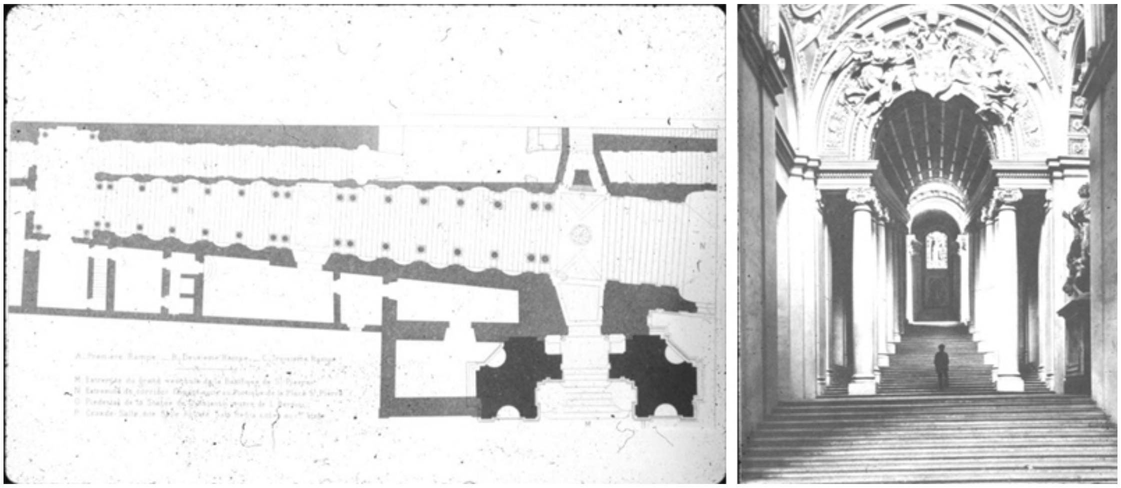


Figure 2.15: Although less dramatically compressed than the Borromini Gallery, Bernini's Scala Regia utilizes a similar warping of the geometry in order to accentuate the perceived depth of the corridor.

Another example of architects using warped geometry to give a greater impression of depth is the apse of the Church of Santa Maria San Satiro in Milan. This example is similar to those by Borromini and Bernini, but was much more dramatically compressed due to the more serious space constraints. Because there was no space in which to build the traditional apse at the head of the church's plan, Bramante designed an interior relief facade that would give the impression of a deep apse (Figure 2.16). The illusion, as was the case in the previous two examples, was based on warping the geometry for viewing from a particular location. The apse at San Satiro is the most compressed of the three examples, but was designed to be viewed from the middle of the central aisle of the church. As the viewer moves closer and off-axis, the apse is revealed even more dramatically to be an illusion (Figure 2.17).

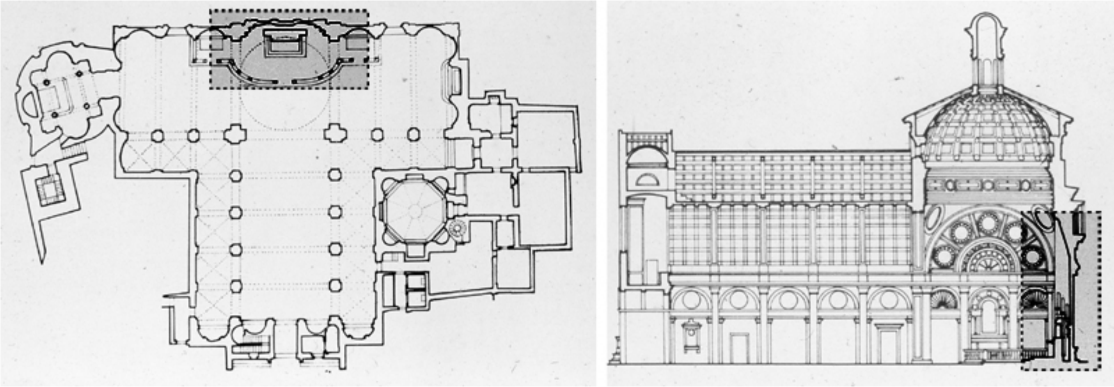


Figure 2.16: Plan *(left)* and section *(right)* views of the Church of Santa Maria San Satiro in Milan show that space constraints forced the designer to dramatically compress the apse of the church. The highlighted areas show that the geometry that gives the impression of a deep apse is actually quite shallow in depth.

Theater designers have long understood the creation of scene geometry that gives an increased sense of depth. They also realized that the illusion of depth is most convincing when viewed from the correct location.[8] An early example of this is the set designed by Scamozzi for the Teatro Olympico in Vicenza (Figure 2.18). When viewed from the correct location, the set gives the impression of seven long streets. A study of the set reveals that the use of perspective geometry extends the illusion of space (Figure 2.19).

The principles behind the illusion used by Borromini, Bernini, Bramante and others are fundamental to the construction of perspective representations and are particularly relevant to the creation of modern computer graphics images. In computer graphics, one creates a perspective image by warping the geometry through a process known as *perspective projection.* After this transformation, the virtual

---

[8] The ideal location for viewing a theater set is usually the center of the eighth row orchestra.

Figure 2.17: Comparison of an on-axis view *(left)* and an off-axis view *(right)* demonstrate the use of a compressed, perspective transformed geometry to give the impression of a deeper apse than could actually have been constructed under the space constraints. When viewed from the nave the illusion is plausible, but an off-axis view reveals that the apse is only a meter or two in depth.

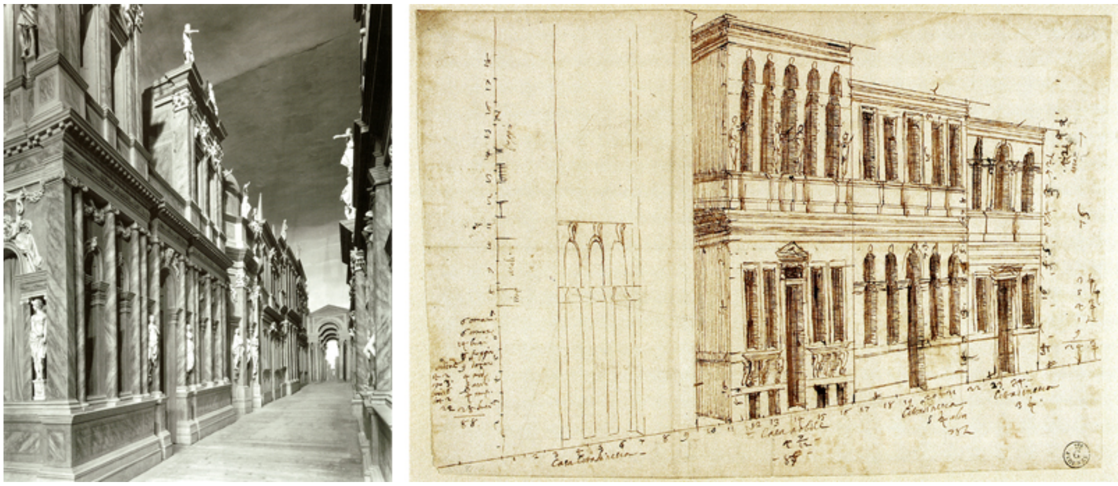


Figure 2.18: The set designed by Scamozzi for the Teatro Olymipco demonstrates the use of perspective geometry in theater set design in order to provide the illusion of increased scene depth.

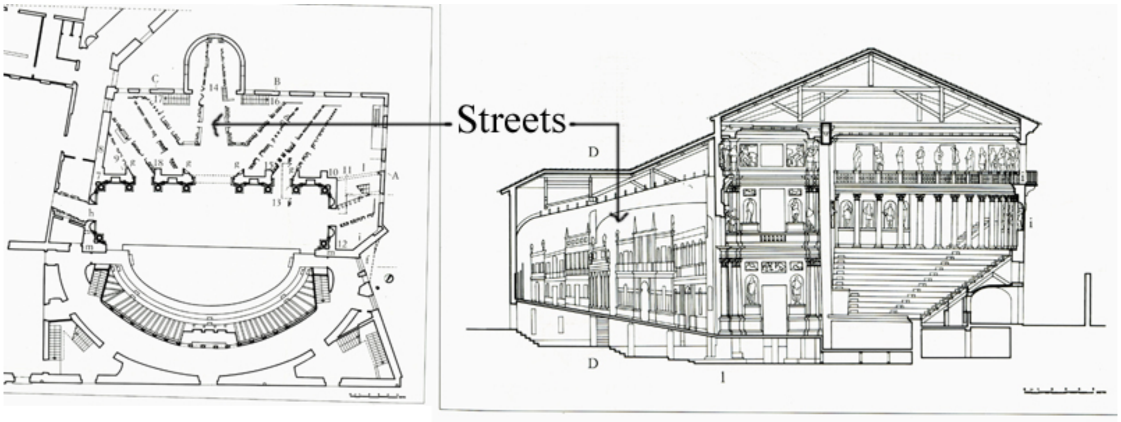


Figure 2.19: Plan (*left*) and section (*right*) views of the Teatro Olympico illustrate the perspective geometry used to create an increased sense of depth. The scene depicts a series of streets that seem longer than they actually are. Note especially the converging lines in the plan and the sloped floor in the section.

scene is still three-dimensional. It does not become a two-dimensional image until it is compressed along the z, or principal, axis. The works of architecture created by Borromini, Bernini and Bramante are essentially equivalent to the scene geometry after transformation by perspective projection. In each case, the resulting warped geometry is more or less compressed along the principal axis depending on the space constraints dictated by the site. The more the geometry was compressed along the principal axis after undergoing a perspective transform, the more severe the distortion realized by a viewer moved from the ideal viewing location. The Scala Regia was compressed only a small amount, requiring the viewer to have to pay close attention in order to notice that the geometry had been warped. The apse at San Satiro, on the other hand, was compressed almost to the point of being a two-dimensional image. This leads an off-axis viewer to clearly recognize the illusion.

We have shown that as the perspectively warped geometry of a scene is compressed along the principal axis, it becomes more important that the viewer's location be known and fixed. The extreme example is a two dimensional perspective image. Following this logic, we contend that in order to more effectively create and display a perspective image, one must take into account the location from which the image is viewed.

2.4.2 Surface Renditions that are Sensitive to Viewer Location

While Borromini, Bernini and Bramante each constructed warped scene geometry, other Renaissance artists painted actual two dimensional perspective images that similarly relied on a fixed viewer location. Many of these, as was the case in the previous examples, used perspective to extend the space of the architecture. In order for the illusion of space to be most convincing, it became necessary that the viewer stand in a specific location.

The necessity of having the viewer of a perspective image in the correct location is made evident by various examples of anamorphic art. An *anamorphic* image is one that seems distorted unless viewed from the proper vantage point.[Sal01] In extreme cases, the distortions are so severe that the image is unrecognizable until viewed from the correct location. In Hans Holbien's Ambassador, for example, one cannot easily *see* the anamorphic portion of the painting as a skull unless viewed from the correct location (Figure 2.20). While some examples of anamorphosis are extreme, these effects are present to varying degrees whenever we view an image from the wrong location. The anamorphic frescos at the apartments of the Gesu in Rome vividly demonstrate this effect(Figures 2.21 and 2.22).

Andrea Pozzo painted two works on the ceiling of the Church of San Ignazio

Figure 2.20: Hans Holbein's *Ambassadors* (*left*) contains a section that, when viewed from a certain vantage point, looks like a skull (*right*). Anamorphic art demonstrates the importance of having the viewer in the correct location.

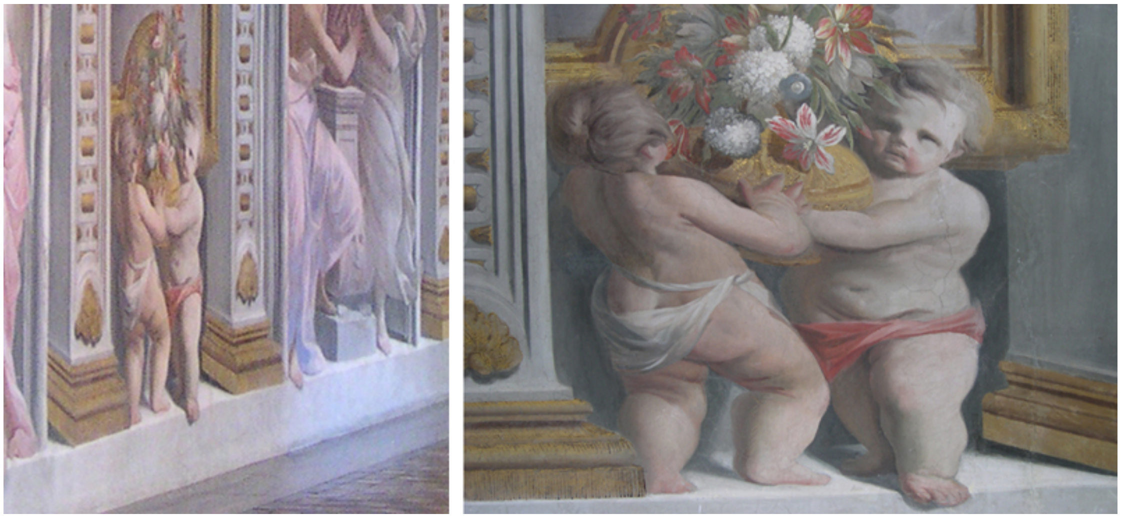


Figure 2.21: Two images of the same fresco at the apartments of the Gesu in Rome. The image on the left shows two cherubs as they appear when viewed from the correct location. The same cherubs, as seen from the wrong vantage point, demonstrate the effects of anamorphism in perspective art.

Figure 2.22: The anamorphic ceiling at the apartments of the Gesu demonstrate an image painted on a curved vault and a flat wall. As seen from the correct location (*left*), the perspective image is one of rectilinear balustrades resting on two arches. From the wrong vantage point (*right*), the architectural elements seem warped.

in Rome that further illustrate the concept of anamorphosis. Above the altar, where it was traditional that a dome be constructed, Pozzo painted a dome on a flat ceiling[9] (Figure 2.23). This illusion holds as long as the viewer is correctly located within the church[10]. [Sal01] The barrel vaulted nave is similarly painted with a scene that requires a privileged viewing location. The scene, entitled *The Transmission of the Divine Spirit*, represents an imaginary architecture of glorious columns ascending heavenward (Figure 2.24). Because the illusion is painted on a

---

[9] As is still the case for those who use computer graphics images for architectural design, Pozzo was taken to task by traditionalists for his virtual depiction of architecture that was "unbuildable". He is quoted as saying,"Some architects disliked my setting advancing columns upon corbels as being a thing not practiced in solid structures; but a certain painter, a friend of mine, removed all their scruples by answering for me, that if at any time the corbels should be so much surcharged with the weight of the columns, as to engender their fall, he was ready to repair the damage at his own cost." [Kem90]

[10] The false dome at San Ignazio has as its ideal viewing location a point within the nave. This coincides with a custom at the time which restricted worshippers to the nave area. Only clergy would have had access to areas of the church from which the illusion would be apparent.

curved surface, the artist invented a system for transforming a planar image to a curved surface. Projections from the ideal viewing location through a planar grid assured that the image would appear the same regardless of the surface on which it was presented(Figure 2.24). Actually, in both cases, since the images are quite distant from the viewer, the illusion holds over a considerable range of viewing positions.

Baldassare Peruzzi, like Pozzo, used a mural in the Salla della Prospettiva (perspective room) in the Villa Farnesina to give the impression of architectural detail where only a flat wall existed (Figure 2.26). Of particular interest in the Salla della Prospettiva was the location of the frescos in relation to the sequence of movement through the room. The intended illusion broke down much more rapidly with lateral movement than it did with movement toward or away from the fresco. This coincided with the way the rooms were arranged so that as one progressed from one room to the next, the fresco's illusory effect was least diminished.

What all these works have in common is that when viewed from the correct location, they give the impression of the intended three-dimensional geometry. But while often quite striking when seen from the correct station point, an off axis view immediately reveals the illusion. The Renaissance artist therefore sought to predict or restrict the location from which their work would be seen.

2.5 Marginal Distortion

Many Renaissance painters used an understanding of station point[11] to accentuate their art without being faithful to a strict construction of linear perspective. In many cases, artists realized that objects in a scene that were located at the fringes

---

[11] The term *station point* refers to the location of the viewer in a scene.

Figure 2.23: This painting by Andrea Pozzo on the ceiling of the Church of San Ignazio in Rome give the impression from a certain vantage point that there is a dome over the altar and not a flat ceiling as is the case.

Figure 2.24: The nave vault at the Church of San Ignazio, *The Transmission of the Divine Spirit*, was painted on a barrel vaulted ceiling. From one viewpoint *(left)* it gives the impression of grand columns extending into the heavens. An off-axis view *(right)* shows the warped columns painted on a curved surface.

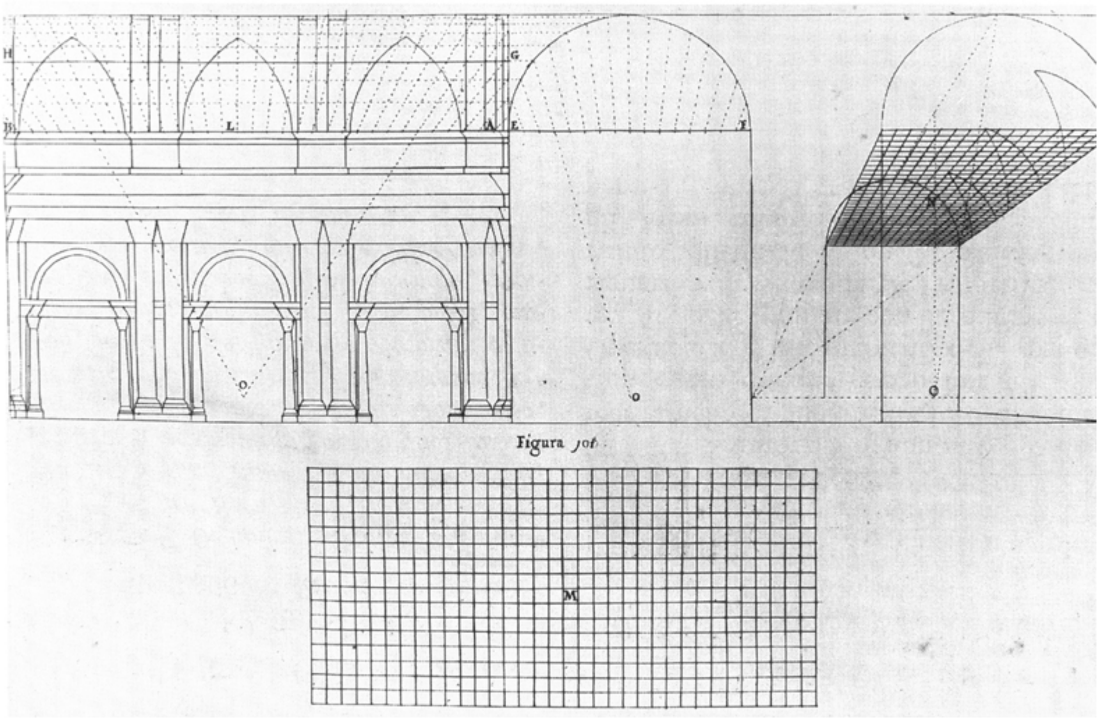


Figure 2.25: This diagram shows how Andrea Pozzo transferred the image of *The Transmission of the Divine Spirit* from a planar drawing to the curved surface of the ceiling. A grid over the image corresponded to a similar grid constructed under the barrel vault. Projections made from the ideal viewing location through the grid to the vault guided the layout of the painting.

Figure 2.26: Peruzzi painted The Salla della Prospettiva to give the impression that there was a column-flanked opening leading out into the countryside. An on-axis view *(left)* is contrasted with a view from the wrong viewing location *(right)*.

of a wide-angle perspective image seemed distorted. The effects that caused this perceptual phenomenon became known as *marginal distortions.* These distortions embody the problem of wide-angle linear perspective.

The creation of the memorial fresco of Sir John Hawkwood in Santa Maria del Fiore by Paolo Uccelo is an example of the early struggle to deal with marginal distortion. Uccelo's motivations for bending the rules of linear perspective illustrate the problem of wide-angle linear perspective and marginal distortion. Upon viewing the fresco, it is immediately evident that there is an inconsistency in the use of perspective (Figure 2.27 *left*). Uccelo painted the horse and rider with a station point that lies level with the rider while the pedestal is painted as if seen from below and slightly to the left. This inconsistency is the result of the incompatible desires to both show the rider's identity while also giving the impression of viewing from a lower, and more respectful, position. After Uccelo finished painting the fresco, the managers of the Opera del Duomo were angered by what they saw as a "dubious tribute" to the deceased general [Kub86]. It was therefore necessary for Uccelo to repaint the section of the fresco that portrayed the mounted rider. One possible reason for the mishap was that Uccelo had painted the rider using the same station point that had been used to depict the pedestal. Because this point was a location far below even the base of the pedestal, the image is essentially "wide", but in the vertical axis (Figure 2.27 *right*). Because Uccelo initially adhered strictly to the rules of linear perspective, the higher portions of the image seemed to be distorted when not viewed from the proper location. This was in contradiction to the rule formalized by Leonardo daVinci which he later followed in his corrected version:

> In drawing from the round the draughtsman should so place himself

> so that the eye of the figure he is drawing is on a level with his own. This should be done with any head he may have to represent from nature because, without exception, the figures or persons you meet in the streets have their eyes on the same level as your own; and if you place them higher or lower you will see that your drawing will not bear resemblance.

Ucello may have initially erred because his depiction of Sir John Hawkwood displayed his head as a seemingly warped ellipse.[12] Leonardo's advice to bend the rules of perspective when depicting spherical objects was intended to avoid such embarrassing situations.

The *central axis of projection* or what Leonardo called the *principle ray* is a line drawn from the station point that perpendicularly intersects the image plane. A spherical object, such as a head, that is not on the central axis of a perspective projection, is presented in the image plane as an ellipse (Figure 2.28). In order to represent spherical objects, painters generally constructed them using their own station points that did not match the rest of the painting. This was done so that each sphere would lie at the center of its own central axis of projection, thus rendering it a circle in the image plane. This concept is critical to the problem of wide-angle linear perspective: The farther a sphere is located from the central axis of projection, the more distorted it will seem if viewed from the wrong location. In the case of Uccelo's mounted rider, the head was far from the central axis of projection. This caused it to look like an ellipse when viewed from the wrong location. In order to avoid this problem, artists painted spheres as circles even at the extreme edges of a wide-angle perspective image.

---

[12] For a modern attempt at solving this ancient problem, see [ZB95].

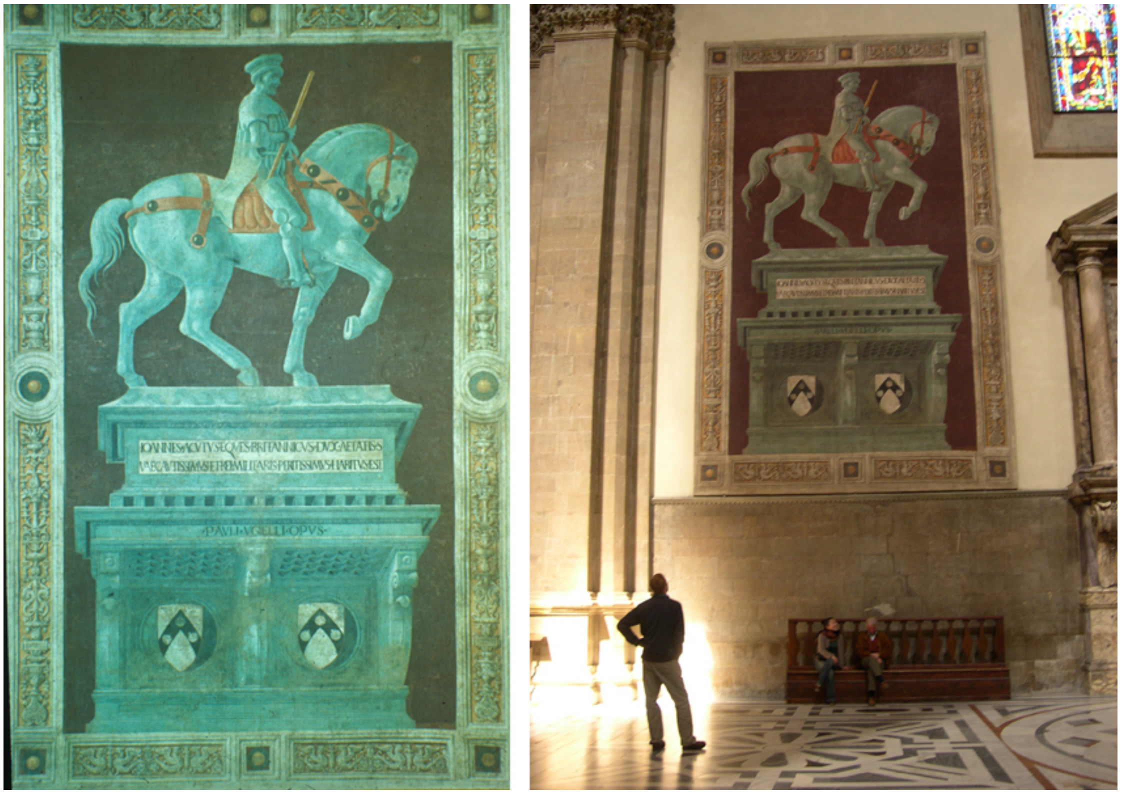


Figure 2.27: This fresco painted by Paolo Uccelo as a memorial to Sir John Hawkwood is an example of a work executed by bending the rules of linear perspective. The horse and rider are constructed using a different station point than was used to depict the pedestal. In the image on the left the pedestal seems to be of poor perspective construction. When viewed from the correct location, as is shown by the viewer in the image on the right, the pedestal is believable as a real three-dimensional geometry.

Another example that illustrates how artists dealt with the adverse effects of marginal distortion is Rafael's *School of Athens* fresco at the Vatican (Figure 2.29). In the extreme right side of the fresco, Zoroaster and Ptolemy hold celestial and terrestrial globes as they converse with one another. These globes, even though they are far from the central axis of projection, are represented as spheres in the image plane. Uccelo, Rafael and others had to abide by an exception to the rule

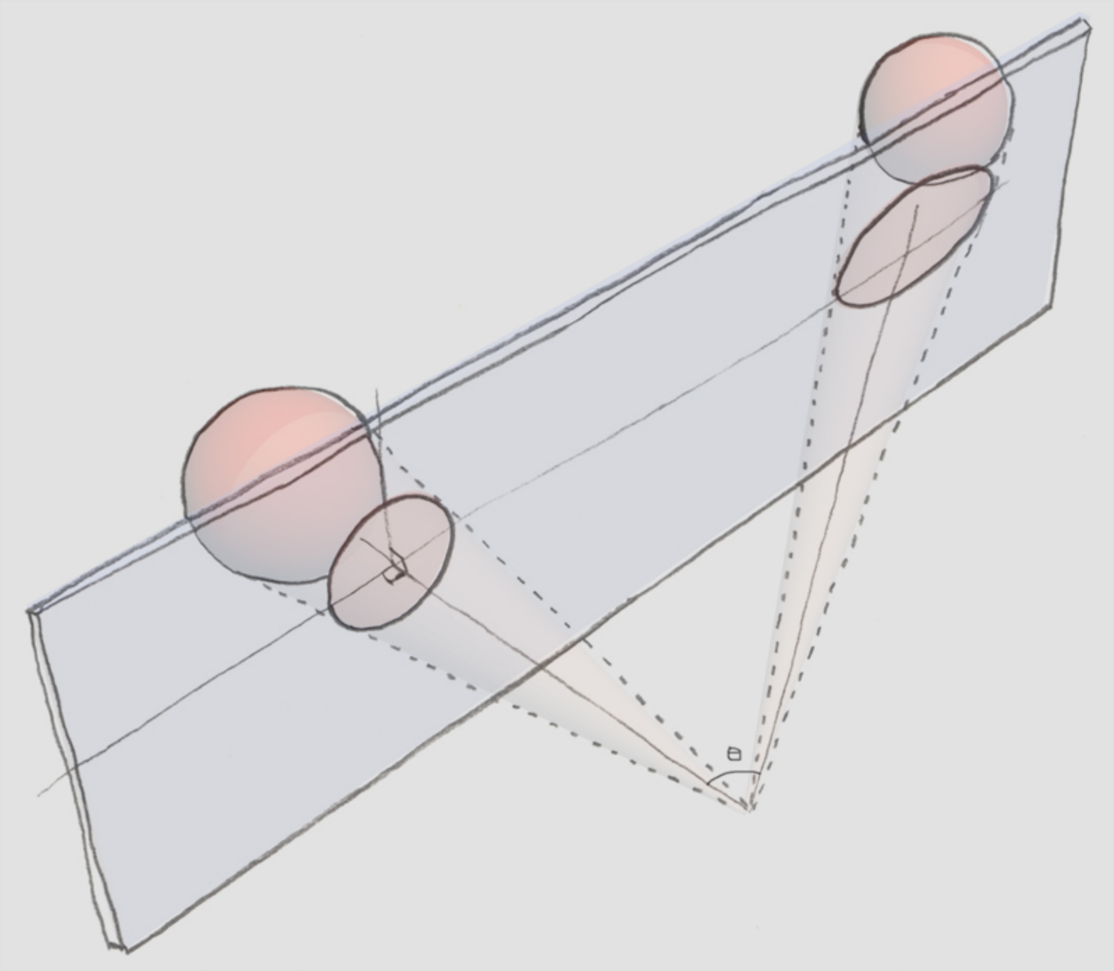


Figure 2.28: The sphere to the left is centered on the principal axis of projection and is represented in the image plane as a circle. The elliptical depiction of the sphere on the right is a result of marginal distortions. As a sphere (or any other object) moves farther toward the edges of a wide-angle perspective image, it becomes increasingly distorted.

Figure 2.29: Rafael's *School of the Athens* shows Zoroaster and Ptolemy (both to the right in the image above) holding globes that are represented as circles in the image plane even though they are off the principal axis of projection. This exception to the rule of linear perspective was one commonly used in Renaissance painting.

of linear perspective in order to present spheres, and especially human heads, in a way that would *look right* to the observer.

A contrasting example is the Statue of David by Michelangelo (Figure 2.30). David's head and brow, while perspectively correct, are anatomically incorrect. The artist created a figure with an abnormally large head and sharp brow because he was certain his statue would always be viewed from below[13]

A comparison of Uccelo's *Memorial to Sir John Hawkwood*, Rafael's *School of Athens* and Michelangelo's *David* illustrate the problem of wide-angle linear perspective detailed by Leonardo daVinci (Figure 2.31). Uccelo and Rafael had to bend the rules of perspective, according to Leonardo's advice, because they could not adequately predict or restrict the location of those that viewed their art. Michelangelo knew that viewers would always see the *David* sculpture from below. Since the David head is not projected onto an image plane, the artist did

---

[13] This is a reasonable assumption when dealing with a statue that is over five meters tall.

Figure 2.30: Michelangelo sculpted the David with a head and brow that, while not anatomically correct, are carved so that they are seen as correct when viewed from below. This is in contrast to the Renaissance artists' rule that spherical objects should be displayed using their own station point.

not have to worry about lateral movement of the viewer revealing the illusion. He could therefore abide by the rules of perspective even though they would result in anatomical incorrectness[14]. As is evident from these three examples, while the viewer's location is known, wide-angle linear perspective images do not present the marginal distortions evident in off-axis viewing. For example, if the viewer's location is known, it becomes possible to display a sphere in the scene as an ellipse in the image without subjecting the viewer to marginal distortions. From other viewing locations the image will seem distorted or anamorphic, but from the correct viewing location proper perception of scale is preserved.

2.5.1 Curvilinear Perspective

One solution that was proposed to the problem of wide angle vision was curvilinear perspective. It ultimately was not widely adopted, but we mention it here because it further illustrates the problem. Linear perspective traces points from a three-dimensional scene toward the station point until they intersect an image plane. Curvilinear perspective traces points from a scene toward the station point until they intersect a curved surface. This method was introduced as early as the Renaissance when Leonardo daVinci described curvilinear perspective, or what he called *natural perspective*, as a solution to the problems of wide-angle linear perspective. He further argued that *artificial perspective*, or the linear perspective used by artists, could only account for wide angle distortions by fixing the eye of the observer [Kem90].

---

[14] As part of a project that scanned many of Michelangelo's sculptures, Mark Levoy created a small scale replica of the David statue. He relates that upon receiving the replica, he was sure there had been an error in its production. The distortions of the head and brow, which had not been noticeable when viewed from below, were apparent when viewed from other vantage points.

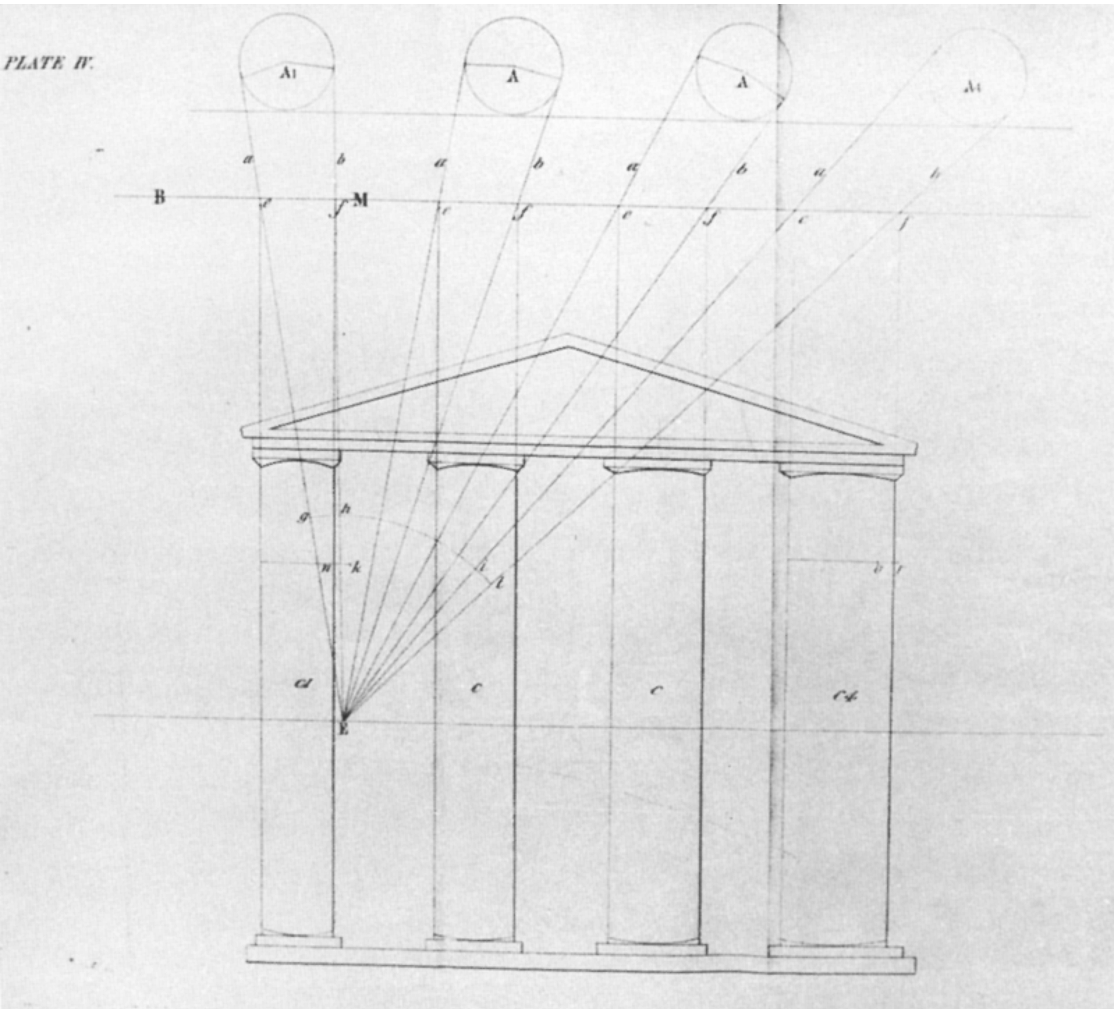


Figure 2.31: The problem of wide-angle linear perspective, as first described by Leonardo daVinci, is here diagrammed by William Herdman. The columns, all of equal diameter, seem to increase in size with greater distance from the central axis of projection.

Curvilinear perspective had a brief revival during the 19th century. William Herdman reintroduced the problem and solution that had been proposed by daVinci. His images, and curvilinear perspective in general, were never adopted (Figure 2.32). Most likely this is because images created using curvilinear perspective that were then displayed on a flat surface seemed distorted to viewers[15].

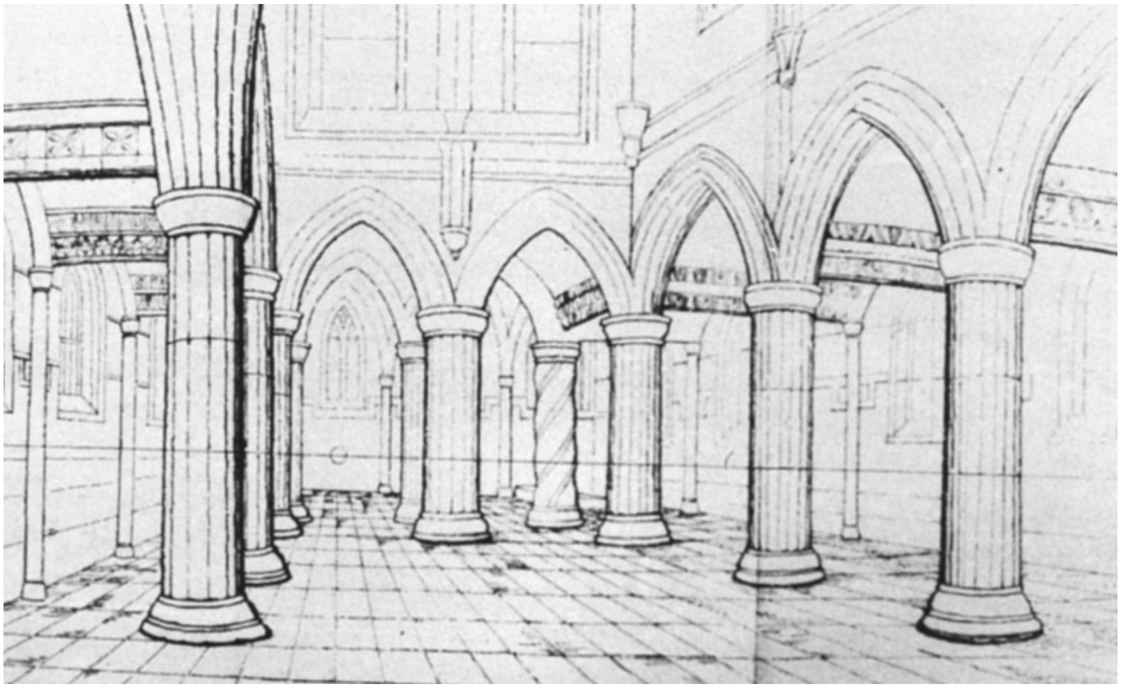


Figure 2.32: An Example by William Herdman of an image constructed using curvilinear perspective.

Another reason curvilinear perspective never became a common method of representation was that the curved lines were difficult to draw. A solution to this problem was the *conform system* devised by Guido Hauk. This method of drawing combined elements of curvilinear and linear perspective so that individual elements

---

[15] This desire to create images using curvilinear perspective coincided with an increased interest in dioramas. These dioramas were relatively successful because they presented the image on a curved surface that resembled that used to create the image.

or sections of the drawing were drawn with their own image plane (Figure 2.33). The separate image planes were arranged around a curved surface. This method is analogous to the more modern application used to stitch individual photographs together to form a panoramic image (Figure 2.34). In this case, each photograph has its own image plane, but together they comprise a perspective image that is roughly curvilinear. Often these panoramas are further stitched to the inside of a sphere, then re-projected to a planar surface in order to be viewed on a planar display[16] Because planar images were the norm, and because it was unrealistic to fix the location of the viewer, curvilinear perspective generally did not see widespread use.

2.6 History Conclusion

In summary, the optical and geometric theories on which the development of the perspective image is based were invented thousands of years ago. However, it was not until the Renaissance that these theories were combined with the practice of art and architecture. This combination led to the creation of perspective images that would give the viewer the impression that they were viewing a real scene. The artists and architects of the Renaissance were sensitive to the location of the viewer, certainly more sensitive than are the projection techniques we use today. They used this recognition of the importance of viewer location to their advantage in creating images that made the observer perceive that they were seeing a *real*

---

[16] $QTVR$ is currently the most popular application of this method. It allows a user to navigate the viewing direction, but fixes the station point. This coincides with our analysis of previous limitations of and wide angle linear perspective. In this case, the field of view is limited and the station point is fixed. For an implementation based on a similar methods of back-projection from a sphere to an image plane, see [TDM01].

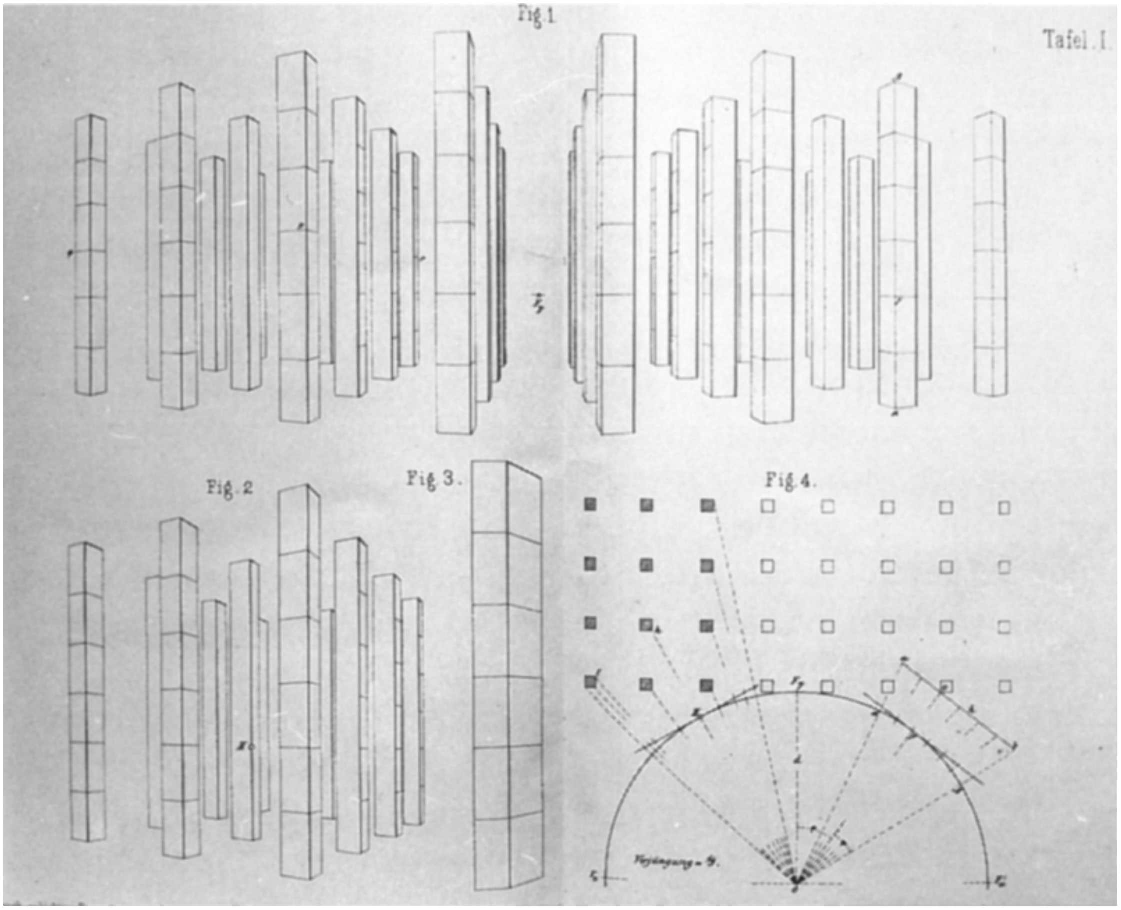


Figure 2.33: Guido Hauk's *Conform System* of perspective representation was a variation on curvilinear perspective. It sought to combine aspects of curvilinear and linear perspective. Discrete image planes trace a curved surface. In this example, each column is drawn using its own image plane that rotates along the circular line. This allows for the construction of small sections of the image using linear perspective, but eliminates major distortions by ascribing to the general principles of curvilinear perspective.

Figure 2.34: A panorama of Schoellkopf field at Cornell University was created by stitching many individual photographs. This is a modern version of curvilinear perspective that conceptually matches the *conform system* devised by Guido Hauk.

scene. They struggled with, but did not quite solve, the problems inherent in the viewing of a wide angle linear perspective image. They knew that the laws of linear perspective were not always appropriate, and should be bent or broken in order to satisfy certain requirements of human perception.[YM04] Yet they also realized that perspective image generation was a powerful tool for showing three-dimensional space.

These are all issues that we still face when using computer graphics images for the display of three-dimensional virtual environments. We commonly use linear perspective as a tool to create images that give the viewer the impression of a real scene, but often the viewer is not in the proper viewing location. Frequently we also try to present an image that fills the field of view by displaying wide-angle views. Unfortunately, viewing a wide-angle perspective image from the incorrect location leads to excessive distortions. These are all issues that have been with us for centuries. Here we have introduced how others have grappled with them and have examined both their solutions and their shortcomings. The concepts discussed in this chapter continue to be relevant to the correct display of perspective images. Later chapters will show how the lessons learned from the history of Renaissance art and architecture can be used to improve the way we create and display computer graphics images.

Chapter 3

The Perspective Image

3.1 Picture Perception

As we observe the world around us, we form a geometric understanding of the space we are in by judging the scale and position of each of the surfaces and objects we see. We assemble surfaces and objects in our mind to create a cohesive understanding of the space we inhabit.[Ren02] Our ability to judge absolute and relative position and scale is based on our perception of depth. Several different visual cues work together to help us judge depth.[1] The relative importance of some of these cues vary with the distance of the scene from the observer (Figure 3.1).

Many of the factors that apply to the perception of depth also apply to the perception of depth in a picture. There are, however, a few differences between our perception of depth in the real world and our perception of the illusion of depth in a two-dimensional perspective image.

Our perception of the space we inhabit is based on the light that enters our eyes. This unique collection of light rays that enter the eye to form a retinal

---

[1] For a more detailed survey of the different visual depth cues, see Appendix A

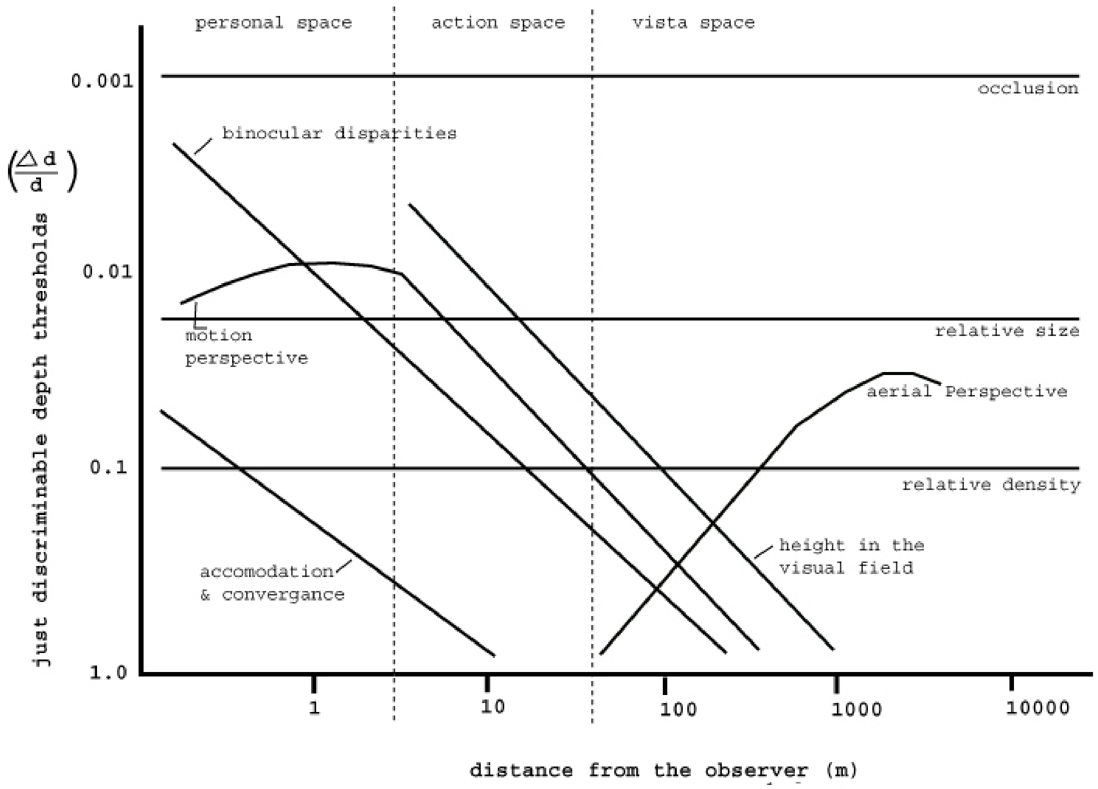


Figure 3.1: This graph illustrates the relative importance of different depth cues as a function of distance from the observer. While some cues such as binocular disparity and convergence are important at near ranges, they become less important as distance increases.

image has been described as a ”bundle of rays.” [FR80] If a perspective image is to adequately replace our perception of the real scene, the ”bundle of rays” that it sends to an observer must replicate those that would reach the observer from the real scene. The light that enters the eye from a perspective image can closely resemble the light that would arrive at the eye from a scene, even to the point that they are indistinguishable under controlled conditions.[Kub86] The fact that this is at least geometrically true is inherent in the way the perspective image is constructed.

Under ideal circumstances, the *bundle of rays*, or *optical flow* that reach an observer from a perspective image would match those that would reach the observer from the scene the image represents. In order for the optical flow to be the same, the retinal image must have the same geometry and the same color intensities as the real (or virtual) world scene. In order to accomplish this, several conditions must be met. In order for the geometry to be the same, we need geometrically correct perspective images. In order for the color intensities to match the real scene, we need accurate material properties that describe the reflection and absorption of light and precise descriptions of the lighting conditions. These conditions must be reproduced using physically correct algorithms for light transport. With all these conditions met, it is possible to create images that are perceptually indistinguishable from the real or virtual scene.

For the past four decades, computer graphics researchers have attempted to create algorithms to accomplish the goal of realistic image synthesis(Figure 3.2). During this time, increased computer power has led to the implementation of accurate lighting, global illumination and recently the physically correct modelling of material properties. Each of these innovations has led to images that, from

Figure 3.2: A major goal of computer graphics research has been to create perspective images that send the same "bundle of rays" to the eye that a real scene would send. This image shows a computer generated image overlaying an image of the real scene. The goal is that the two be perceptually indistinguishable.

a lighting standpoint, can more closely match a real scene. From a geometrical standpoint, the images are based on the assumption that planar affine perspective transforms yield correct geometrical images. This assumption is valid with one important caveat: Implicit in the standard procedures is that the viewer is in the correct viewing position. This position must be either the same as, or congruent to, the camera position in the virtual scene. In practice, this specific positioning of the viewer of a perspective image is rarely upheld.

In order for a perspective image to provide a geometrically accurate representation of a scene, the viewer must be correctly located. An example of this phenomenon is evident in the creation and display of photographic images. While some images seem to have an extreme *telephoto* or *wide-angle* effect, the real problem is that we might be viewing them from the wrong vantage point. Correct viewing location for a photographic image is proportional to the focal length and the rate of enlargement (Figure 3.3). If the photographer uses a skew camera, the location for optimal viewing of the print can be calculated using the amount of lens offset, the focal length, and the rate of enlargement (Figure 3.4). These types of images are used quite often in architectural photography because they preserve a sense of relative scale when viewing a major facade (Figure 3.5). A traditional camera with a film plane that is fixed perpendicularly to the centerline of the lens creates images that seem distorted. This is because vertical parallel lines in the scene converge. The skew camera eliminates these distortions by positioning the film plane vertically even when the camera is pointed upward.

If the viewer is not in the correct location, the geometry that they see represents some scene other than the one used to create the image (Figure 3.6). The same image viewed from two different locations sends two dramatically different ”bundles of rays” to the observers. The comparison of the optical flow reaching two observers from the same perspective image of a simple scene demonstrates this effect (Figure 3.7).

We have shown that the location from which an observer views a perspective image has an effect on the ”bundle of rays” that arrives from the image to the observer's vantage point. Yet an observer can view a perspective image from various points and still get a sense for the three-dimensional space the image represents. As was noted by Farber and Rosinski, ”We perceive a pictorial representation of space veridically, even when the geometric projection to the eye is greatly distorted.” [FR80] We have learned to compensate for the viewing of perspective images by reconstructing the correct vantage point in our minds and placing ourselves there.[2]

---

[2] This concept has interesting implications in studying the historical examples in Chapter 2. One could argue that Renaissance painters were trying to ”place” the viewer of a painted scene in a particular relationship to the painting and therefore

The ability to compensate for incorrect viewing location has its limitations. These limitations are made most evident by the effects of anamorphosis and marginal distortion discussed in Chapter 2.

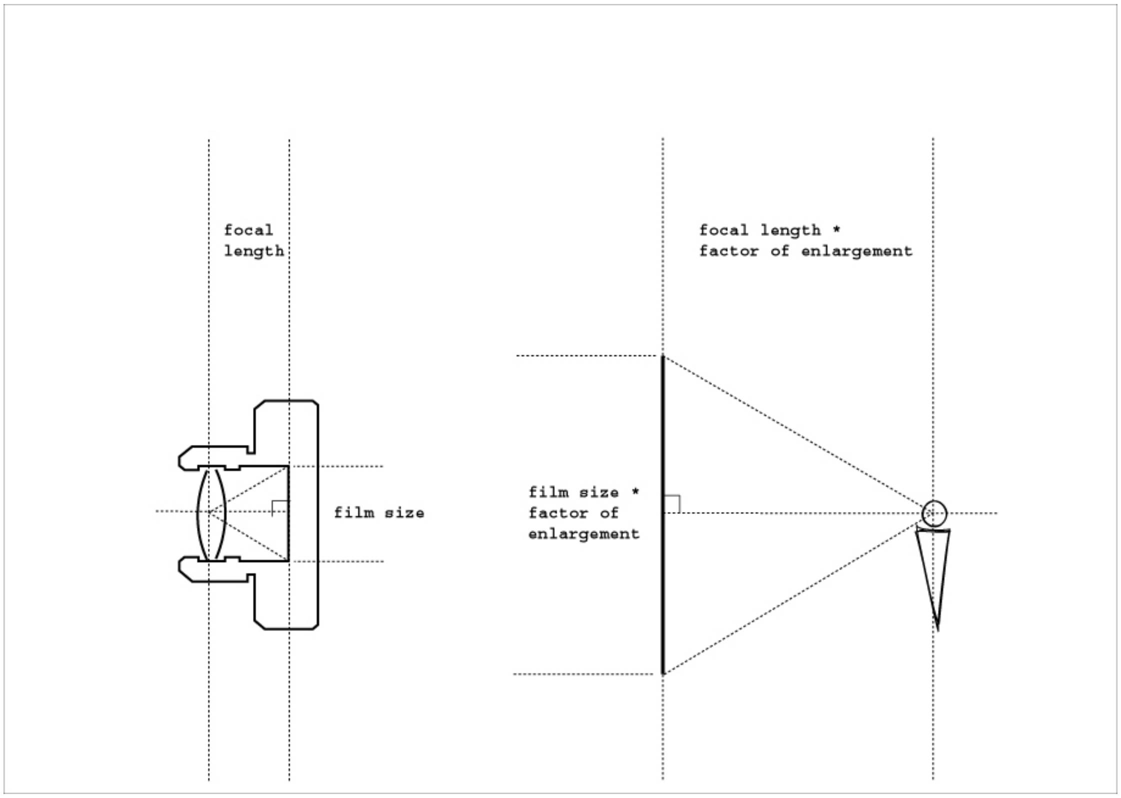


Figure 3.3: A diagram of a camera's optics (*left*) along with the elevation in real space (*right*) of the conditions for viewing the resultant image. In order to find the ideal location from which to view a photographic image, follow a line drawn perpendicularly from the center of the image a distance equal to the focal length of the lens multiplied by the factor of enlargement of the print. An image which at first might appear to have strong wide angle distortions might actually be the result of viewing the image from too far away. The reverse is true for an image that seems to have a strong "telephoto effect."

---

the subject. i.e. Masaccio's Trinity and Ucello's Monument to Sir John Hawkwood.

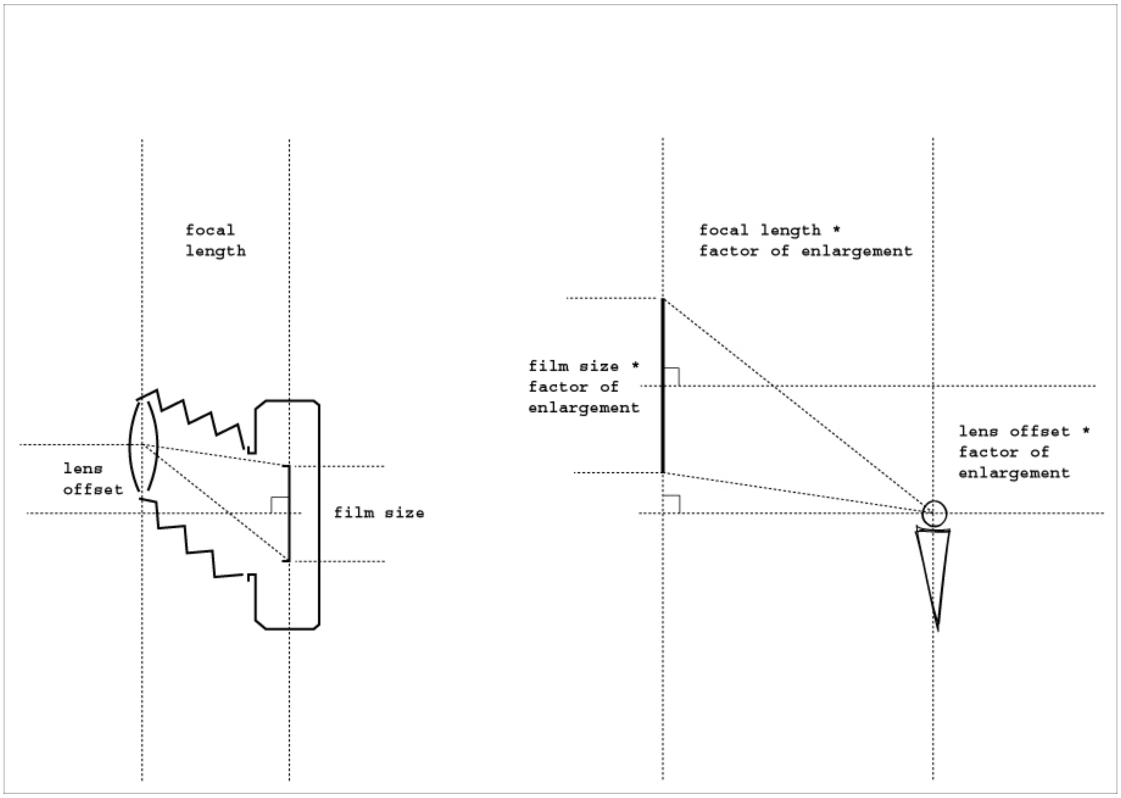


Figure 3.4: The method for finding the location from which to view a photographic image that was taken using a skew (or view) camera is similar to that used to analyze a rigid body camera. The only addition is that of a lense offset multiplied by the factor of enlargement for the print.

Figure 3.5: A comparison of two types of cameras typically used for architectural photography demonstrate the difference between a standard viewing camera (top) and a skew camera (bottom). The skew camera is often considered ideal for photographing architecture because the image plane can be aligned with a major facade independent of the direction of view - which in this case is angled upward in order to encompass the entire scene.

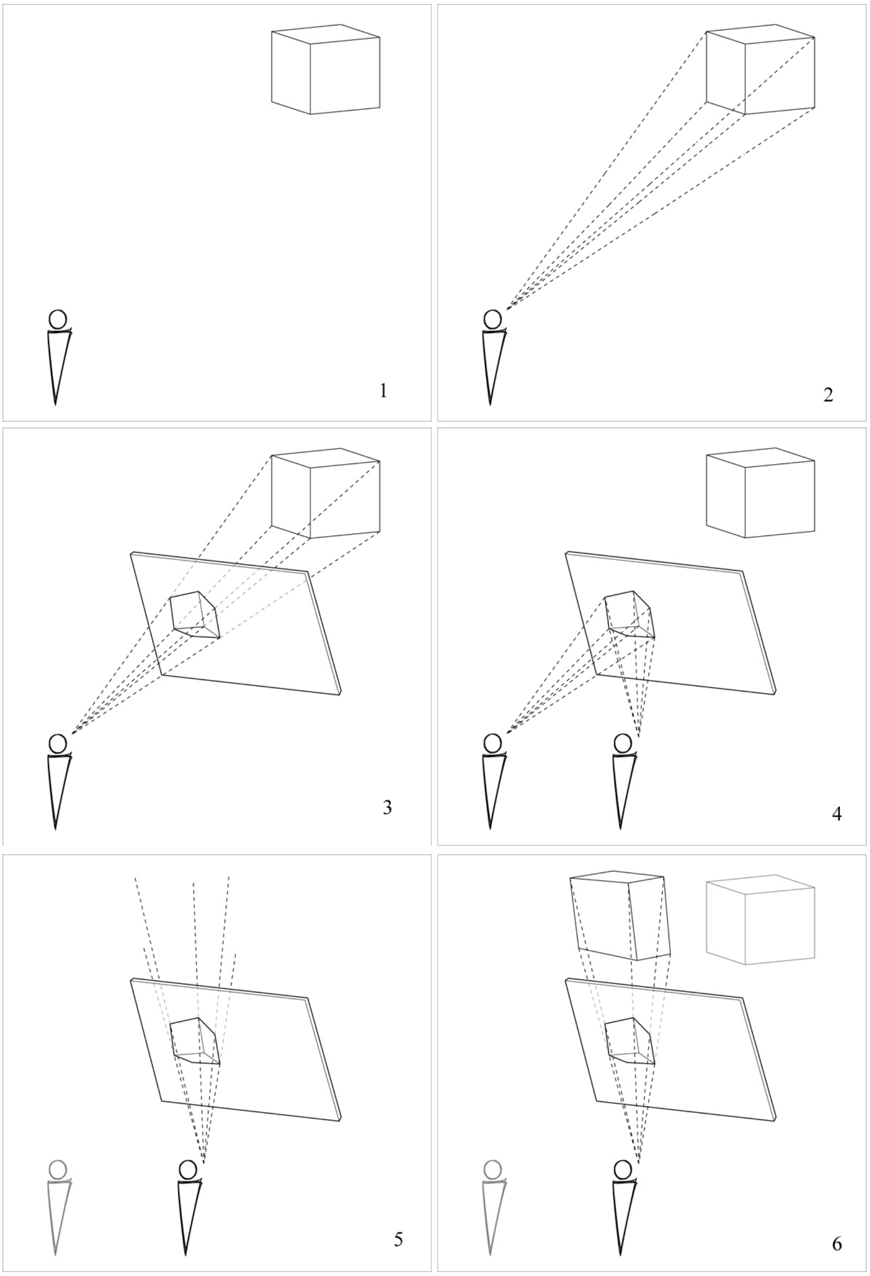
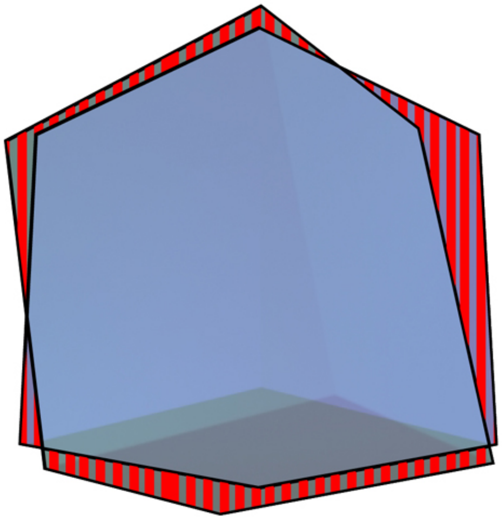


Figure 3.6: Diagrams showing how a "bundle of rays" can send the same information to the observer of an image that will be sent from the real scene. 1)an observer and a simple scene 2)the bundle of rays that reach the observer from the scene lead to the observer's perception of the scene 3)an image can send the same bundle of rays 4)two observers looking at the same image 5)the second observer 6)the bundle of rays that reach the second observer give the impression of some other geometry.



..

Figure 3.7: A perspective image of a cube viewed from two different locations sends two dramatically different "bundles of rays" to the observers. Here we show one viewed over the other in order to compare the optical flow that reaches both observers. This comparison is an example of the diagram in panel four of Figure 3.6.

Chapter 4

View Dependent Perspective Images

Chapter 4 outlines the details of a system that shows view dependent perspective images to an observer located anywhere within a viewing environment. This particular application is an example of how one might develop computer graphics tools that better take into consideration the relationship between viewer, image and scene. In general, a perspective image that takes into account the viewer's location will not suffer the ill effects of marginal distortion. These observations allow us to generate perspective images of virtual three dimensional environments which look correct from any viewers position. These view dependent perspective images provide designers with a better sense of space and scale.

Additionally, these techniques can be used for improved renditions of wide-angle views. When using wide-angle lenses, distortions always occur near the boundaries of the image. The same phenomenon holds true when applying linear perspective transforms to wide-angle views. By segmenting the image into a set of smaller narrower angle views the distortions can be minimized. This makes possible the display of wide-angle perspective images without the perception of scalar distortions. We have created these images by tiling multiple displays to create a single wide-angle image. By providing perceptually correct wide-angle views, we can improve the designer's sense of space and scale.

4.1 Introduction to the Problem

As we have previously discussed and illustrated, designers would like to be able to get a better sense of space and scale when viewing digital images. One way to do this is to fill more of their field of view by using wide-angle displays. We have shown that the extremes of a wide-angle image become increasingly anamorphic, and must be viewed from the correct location. This presents the need to know a viewer's location in real time without encumbering tracking apparatus. We have further found that in order to take advantage of parallel processing, and thus increase image resolution, we can segment the image, render in parallel, and then tile the displays. [RWC^{+}98] [YGH^{+}01] This approach necessitates that a tracking device pass its location data to multiple rendering machines. For flexibility and scalability these displays should have the ability to be located and oriented in any configuration.

4.2 User Location Tracking

We have shown that it is important to view a perspective image from a specific location. Ideally, every perspective image would be created based on the known location of its viewer. Many technologies now exist for finding a viewer's three-dimensional location. Several factors influence the capabilities of any location tracking system. The particular needs of any application determine which technology is used.

The accuracy of a tracking system is referred to as its *resolution.* A tracker with a higher resolution will be able to more accurately distinguish between two locations that are close together. For example, a system with 2mm of resolution would give accurate location data, but have an error tolerance of 2mm. *Latency* refers to the lag that occurs between an update in the location data and an applications retrieval and use of the data. A tracking system with low latency would cause the user of a graphics system to have a lagging sense of "swimming." [Bro99] The *refresh rate* indicates the rate at which updated location data is made available. A refresh rate of 30 Hz, for example, would provide updated location data at a rate of 30 times per second. 30 Hz is a common benchmark used by those interested in implementing an application that operates at interactive rates. The effective *range* indicates the tracked area, usually in reference to some base unit. If a tracking system has *line of sight* issues, this means that the ability to track a target will diminish or disappear if the target is occluded.[1] A significant factor in implementing a tracking system that will be widely used is *cost.* Any effort to make the system widely accessible to design communities would have to take this into consideration. Another factor that relates closely to accessibility is the ease of use. The robustness of the system itself makes it useable by designers without full-time attention from a technical specialist. Additionally, the level to which the tracking apparatus interferes with the user's behavior affects its adoption and widespread use by designers. For this reason, a tracking system that is contact free provides less distractions for the user. [TL99][OM04] There are many options

---

[1] Occlusion, as seen by a base unit, can be either visual or otherwise depending on the method of tracking.

available for the acquiring of user location data.[2] We have presented the criteria by which one can evaluate these options.

4.3 Camera Based Tracking

Many options exist for tracking a viewer's location in space (Appendix B). We have chosen to use a camera based system because it combined the necessary resolution with an adequate refresh rate. Ultra-Wideband tracking technology, because of its increased resolution and range, would have been an alternative choice, but because the technology is still fairly new, even a small tracking system was too expensive.[FRB03] Thus, we implemented a two camera infrared tracking system that provides user location data in real time for a cost of under two hundred dollars.

Our tracking system was based on concepts of epipolar geometry used in the computer vision community. We compared views of a target as seen by a stereo pair of cameras in order to find the target's location in three-dimensional space (Figure 4.1). We used two standard webcams[3] that capture images at 30 frames per second with a resolution of 640x480 pixels, each tracking the infrared reflection from a retro-reflective target.[CKKP] Because most webcams are set up to filter infrared light, it was necessary to remove the infrared filter housed between the lens and the charge coupled device (CCD). Results were improved by replacing the infrared filter with two layers of exposed 35mm film negative. This blocked visible light while allowing infrared light to pass through.

Each camera was paired with an array of 36 infrared light emitting diodes (LEDs). The LEDs were arranged and mounted so that they had a targeted area

---

[2] For a more comprehensive evaluation of the methods most commonly used for location tracking see Appendix B. [WF02] provides another helpful survey.

[3] QuickCam Pro 4000

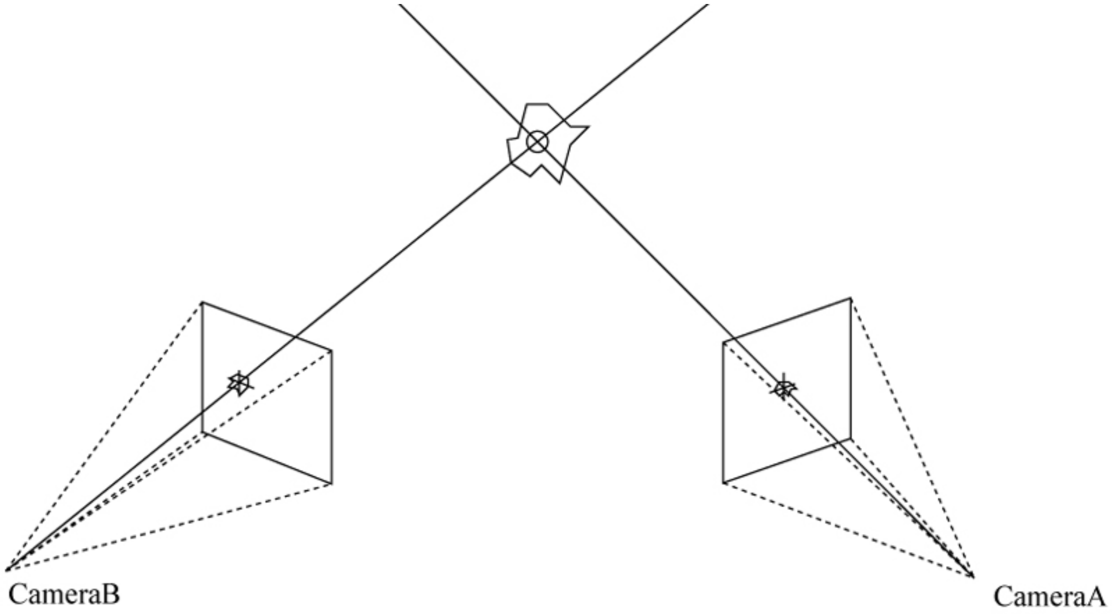


Figure 4.1: Two cameras together can locate a target in three-dimensional space. Cameras A and B each identify the target in image space and use this information to specify a direction along which the target must lie. The point at which the two rays intersect is the target location.

that was equal to the field of view of the cameras (Figure 4.2). We used retro-reflective tape as a target. Because the tape reflects light back toward its source thousands of times more than does an ordinary surface, we were able to use an intensity cutoff in each camera's image to include pixels that fell within the target area. Finding the center of mass for all included pixels gave us an image space location of the target's center (Figure 4.3).

Figure 4.2: An image of one of the two cameras used to track users within a specified area. Each camera was paired with an array of infrared LED's that was aimed and targeted to match the camera's field of view.

After finding the target in image space for each camera, we calculated the vector direction in camera space between the camera and the target.[4] After rotating the vector from the camera's basis to world co-ordinates we created a ray that began at each camera and went in the direction of the target. The intersection of these two rays yields the target location. Because error tolerances dictated that the rays

---

[4] Because we use a left-handed orthonormal basis, a direction of 0,0,1 in camera space would indicate a target that lies along the central viewing axis of the camera, or at the camera's target.

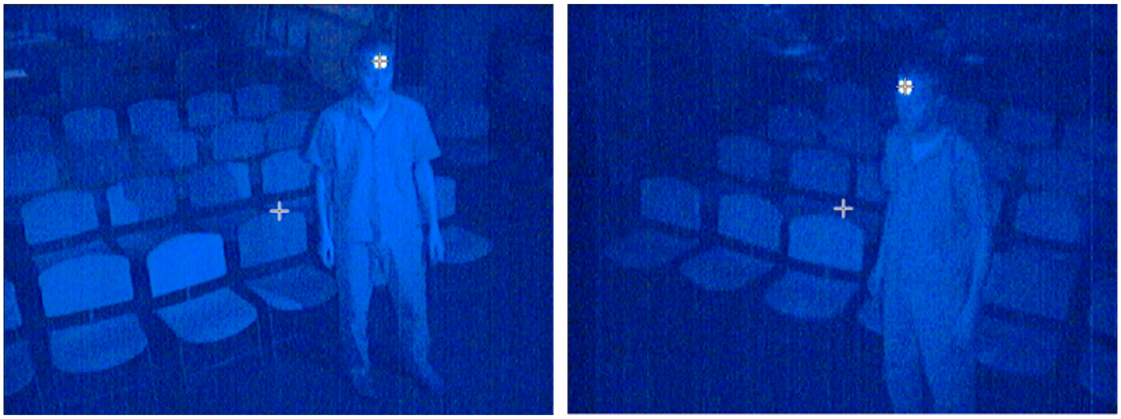


Figure 4.3: The image of the tracked area as seem from camera A (left) and camera B (right). The retroreflective tape attached to the target is so much brighter than the surrounding environment because each camera was paired with an infrared source. Additional modifications to the camera made it more sensitive to infrared light and less sensitive to visible light.

would seldom exactly intersect, we used as the target location the midpoint of the shortest line that connected the two rays. This system provided the location of the viewer in real time with adequate accuracy. Placing the cameras closer together increased the size of the tracked area, but decreased the accuracy with which depth was measured. We found that lateral accuracy was relatively more important than was accuracy in depth. When combined with the location of the display this tracking system allows for the construction of an accurate perspective image that is unique given viewer and display location.

4.4 Skew Camera

A camera in which the picture plane is not perpendicular to the line of sight is commonly known as a *skew camera* (Figure 4.4). The line of sight is a line that connects the camera to the center of the image. Another way to think of a skew camera is to consider a setup in which the image plane is not centered around the central axis of projection. Skew cameras have a relatively long history, especially in architectural photography. Their use allows for more accurate perception of space and scale in photographs of architectural space. Their use in computer graphics viewing however has been quite limited. This section describes the implementation of a skew camera for computer graphics viewing (Figure 4.5). The skew camera marix allows one to view a scene from any location through an arbitrary picture plane. This permits flexibility and scalability when using the skew camera as a base for the creation of applications that will aide the architectural designer.

We first describe the steps used for transforming a scene using a standard camera in order to compare them to those used by a skew camera.[5] The standard camera most often used for viewing in computer graphics uses a *perspective projection matrix* to perform a *perspective transform* The matrix that performs this operation is defined by Equation 4.1 [Shi02].

$$M_{projection} = \begin{bmatrix} \frac{2|n|}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2|n|}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & \frac{|n|+|f|}{|n|-|f|} & \frac{2|f||n|}{|n|-|f|} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{4.1}$$

Each point in a scene is multiplied by this perspective projection matrix. The result is a warped geometry that, when compressed in the z-axis, yields a perspective image. The projection matrix is defined by using the values found in Equation

---

[5] We only present the standard sequence of transforming a scene to the extent needed to contrast it with that used to transform a scene using a skew camera. For a more comprehensive introduction to perspective projection see a computer graphics textbook such as [Shi02].

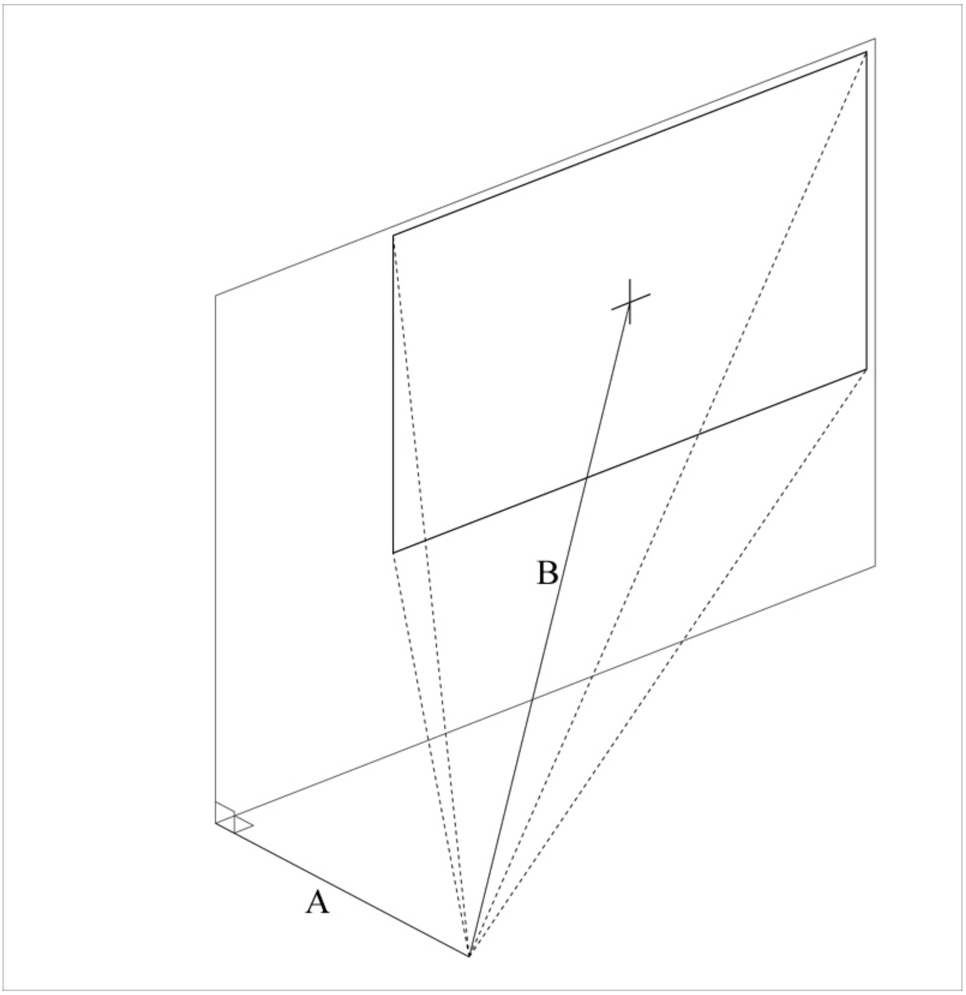
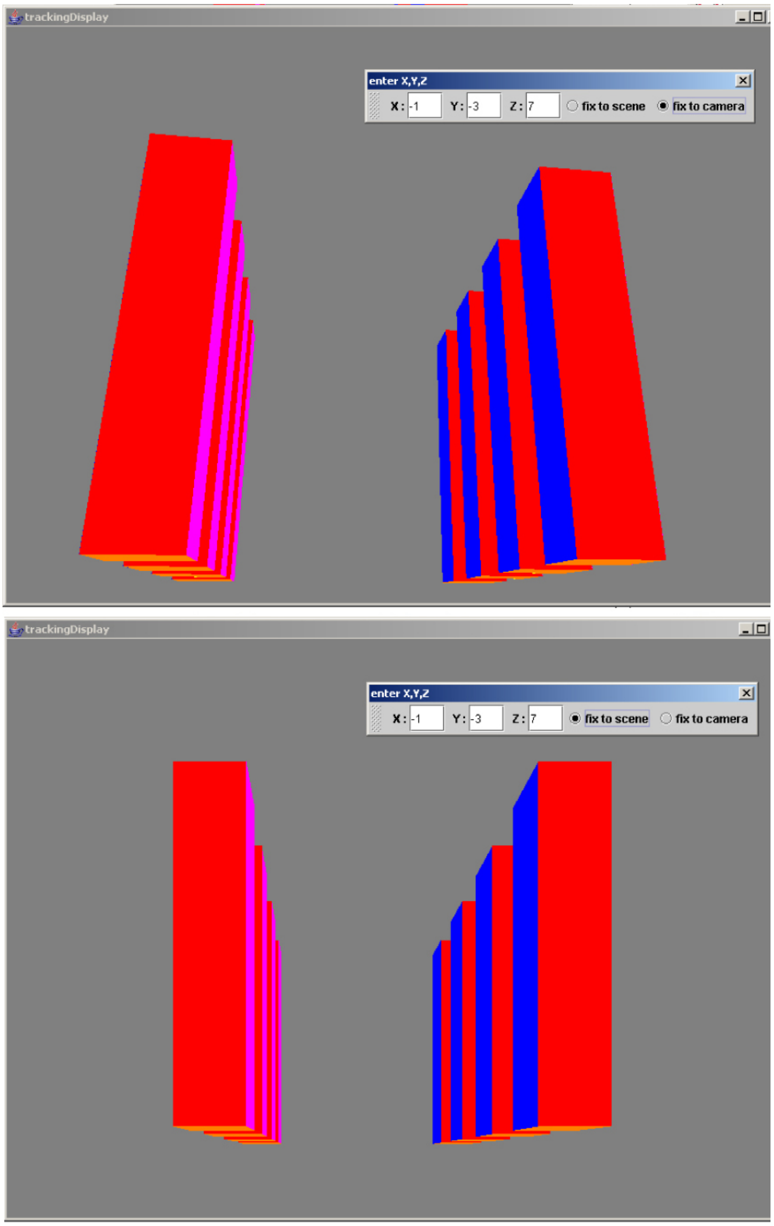


Figure 4.4: A skew camera is one in which the image plane is not centered around the central axis of projection (line A). Another way to think of the skew camera is to draw a *line of sight* that connects the camera to the center of the image plane (line B). If the line of sight is not perpendicular to the image plane, the camera is a skew camera.


trackingDisplay

enter X,Y,Z

X: -1 Y: -3 Z: 7 ○ fix to scene ● fix to camera

trackingDisplay

enter X,Y,Z

X: -1 Y: -3 Z: 7 ● fix to scene ○ fix to camera

Figure 4.5: The top image shows a scene using a standard perspective camera. The bottom image shows the same scene using a skew camera. The skew camera is often used in architectural photography because it gives a better sense of relative scale, but is seldom used by designers in a computer graphics setting.

4.2 (Figure 4.6).

$$
\begin{aligned}
n &= nearplane = k_1 \\
f &= farplane = k_2 \\
l &= left = -right \\
r &= right = -left \\
b &= bottom = -top \\
t &= top = -bottom
\end{aligned}
\tag{4.2}
$$

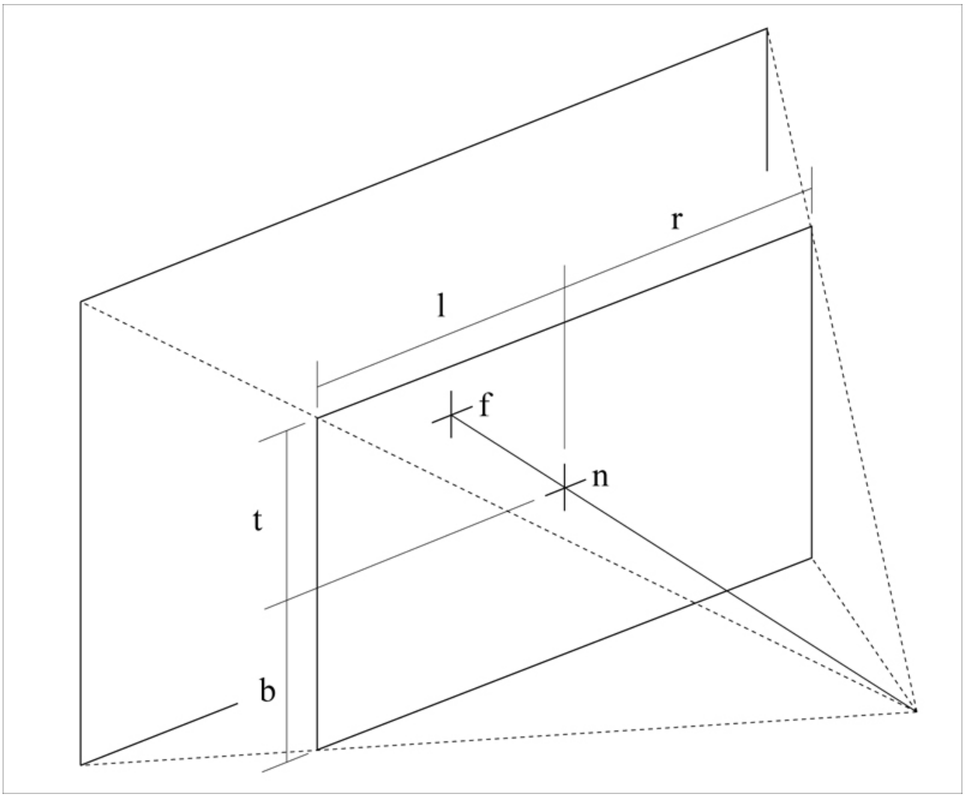


Figure 4.6: The standard viewing frustum as described in Equations 4.1 and 4.2. Note that distances are measured from the viewer, or camera, position.

Most applications define a camera by its aspect ratio and field of view. For this reason, the terms for top, bottom, left and right used in the projection matrix are derived using the resultant image's size and the field of view of the camera.

The perspective transform assumes that the camera is at the origin and that the image plane is parallel to the xy plane. In the case of a standard viewing camera, this means that the camera faces in the $-z$ direction. As this is very rare in any scene, a preliminary step requires that the scene be translated and rotated so that the camera is at the origin and facing in the $-z$ direction (Figure 4.7). This is done using the camera's location and target. After this step is complete, the projection matrix transforms each point within the frustum so that it lies within the canonical view volume (defined by Equation 4.3). The resulting distorted geometry that is found within this volume is used to create the perspective image.

$$\begin{aligned} x &\in [-1, 1] \\ y &\in [-1, 1] \\ z &\in [0, 1] \end{aligned} \tag{4.3}$$

Implementation of a skew camera requires an additional step beyond those used for the normal viewing described above. That step involves the skewing of the scene and the camera so that the camera's frustum resembles that of a traditional symmetrical viewing camera (Figure 4.8). This step occurs after the translation of the eye to the origin, and before the standard perspective transformation. In both cases, all the points that lie within the camera's frustum are transformed so that they lie within the canonical view volume (Equation 4.3). The whole process is accomplished for a skew camera by transforming each point using the same matrix used for the standard camera (Equation 4.1). However, the methods used to place the camera and fill the matrix for a skew camera differ from those used for standard viewing.

The first difference lies in the rigid translation and rotations that are used to place the camera prior to the execution of the perspective projection (Figure 4.8). As in the standard camera, the scene is translated so that the camera lies at the origin and faces in the $-z$ direction. By definition, the skew camera's *line of sight* is not the same as the central axis of projection (Figure 4.4). When rotating the camera, it is necessary that the central axis of projection, and not the *line of sight* be aligned with the z axis. In order to do this, the camera is rotated so the image plane is parallel with the xy plane.

The second difference lies in the method used to fill in the matrix based on the camera's parameters. By definition, the left boundary of a skew camera's image plane (l) is not necessarily equal to the absolute value of the right boundary (r). The same is true for the top (t) and bottom (b) boundaries. This is because the skewed picture plane is not symmetrically aligned around the central axis of projection. Because $l \neq -r$ and/or $b \neq -t$, the $\frac{r+l}{r-l}$ and $\frac{t+b}{t-b}$ terms in the perspective matrix to not disappear to zero as they do under normal viewing. These are the two terms that skew the scene as part of the perspective projection. In this context, the skew camera can be conceived of as a portion of a larger wide-angle image (Figure 4.8). If the boundaries of a camera's image plane are analogous to the boundaries of a display, then a skew camera represents what should be displayed when the viewer is not seated directly in front of the image. The width (w) and height (h) of an image plane that is in the xy plane and centered at the origin, along with the camera location (x,y,z) can be used to find the necessary values to fill in the perspective transformation matrix for a skew camera using Equation 4.4.

$$\begin{aligned}
n &= nearplane = k_1 \\
f &= farplane = k_2 \\
l &= left = -(x + \tfrac{w}{2}) \\
r &= right = \tfrac{w}{2} - x \\
b &= bottom = -(y + \tfrac{h}{2}) \\
t &= top = \tfrac{h}{2} - x
\end{aligned} \tag{4.4}$$

The use of skew camera transformations will form the foundation for the creation of applications in computer graphics that provide an increased ability by designers to judge space and scale in a digital environment. Its flexibility allows for viewing a scene from an arbitrary location through an image plane of arbitrary location, size and orientation. Thus, if the viewer is not in precisely the right location to view the projected picture, the correct perspective image can be computed for the viewer's known position. As these locations are known and tracked, the relationship between the viewer, scene and image will assume the same sensibility inherent in the examples of Renaissance art and architecture described in Chapter 2. But unlike the historical examples cited, location tracking and computer graphics technology will allow the image to dynamically respond to changes in viewer and/or picture location.

From a perception psychology point of view, the use of a skew camera that is based on user location tracking provides for more geometrically accurate perspective images. This is because the optical flow that reaches an observer can be the same as that which would arrive from the real scene (see Chapter 3). This changes the computer graphics paradigm from that of looking at a picture to that of looking through a window.

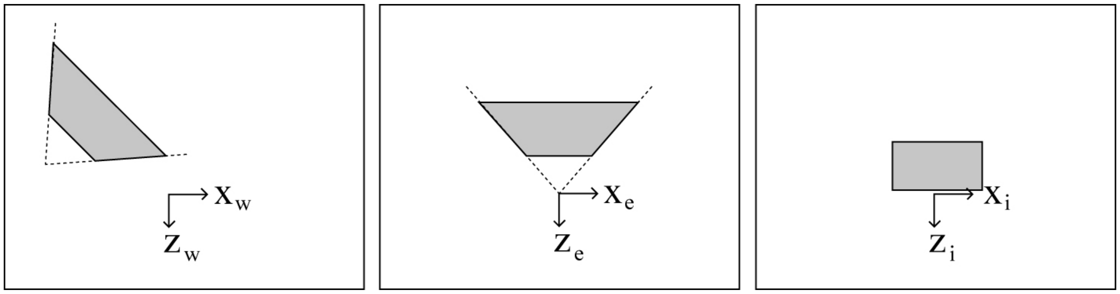


Figure 4.7: These diagrams illustrate the sequence of transformations that occur in order to view a scene using a standard perspective camera. The first step is a rigid transform that moves the scene so that the camera is at the origin and facing in the -z direction. The second step perspectively transforms all the points that lie within the viewing frustum so that they lie within the canonical view volume(from -1 to 1 in x and y, from 0 to -1 in z). These steps take the scene from world co-ordinates(w) to eye space(e)co-ordinates to image space (i) co-ordinates.

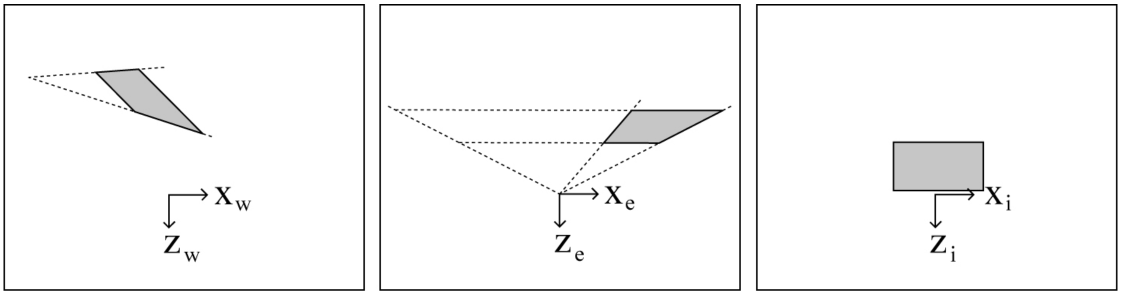


Figure 4.8: These diagrams illustrate the sequence of transformations that occur in order to view a scene using a skew camera. The first step is a rigid transform that moves the scene so that the camera is at the origin and the image plane is parallel to the xy plane. The second step skews the entire scene and applies a perspective transform so that the entire frustum lies within the canonical view volume. The center panel shows the skew camera as part of a larger wide-angle image. Only the shaded portion is transformed to image space using the skew perspective matrix.

4.5 Tiling Displays and Networking

To create a single wide-angle perspective image, we implemented a networking system that allows for the use of multiple displays (Figure 4.9). Each display utilized the skew camera described above, and takes input from a camera based tracking system. As with the single projection system, the tracking system acquires a frame from each of the two tracking cameras[6]. Using the known resolution and field of view for the cameras, the tracking system then calculates a vector that describes the target direction in each camera's basis. These vectors are then placed on a port for retrieval by rendering machines. The viewer location derived from these vectors is the same for each projected image.

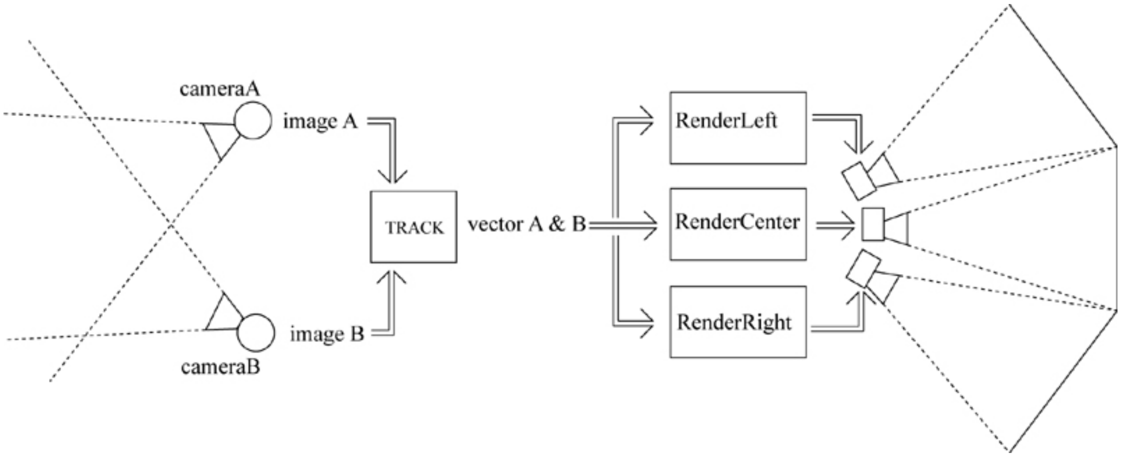


Figure 4.9: This diagram illustrates the network that was implemented to allow for the tiling of more than one display to create a single wide-angle perspective image.

The number of displays is scalable. We tested the setup using three tiled displays in the Cornell Program of Computer Graphics conference room (Figure 4.10). Each of three rendering machines follows the same process in parallel. After re-

---

[6] Because the cameras were identical in make and model, differentiating their IDs required that they be entered into the tracking system in a known sequence.

Figure 4.10: Three displays tile to form a single wide angle perspective image. The image is updated based on the viewer's location. Note that in this figure, the width of the equally sized columns look wrong since the real camera used to take this picture is in the wrong location. However, from the viewer's point of view the geometry appears correct (Figure 4.11).

trieving the two vectors from the tracking machine, calculations are made using the camera's location, target direction and field of view to find the viewer location (xyz). The viewer's location, along with each display's size, location and orientation defines a unique skew camera for each display. We render using this skew camera, look for new location data and repeat the process. In our setup, the wide horizontal field of view of the three tiled displays leads to marginal distortions at the edges of the image. These distortions are evident when viewed from the wrong viewpoint, but look correct when viewed from the correct location (Figure 4.11).

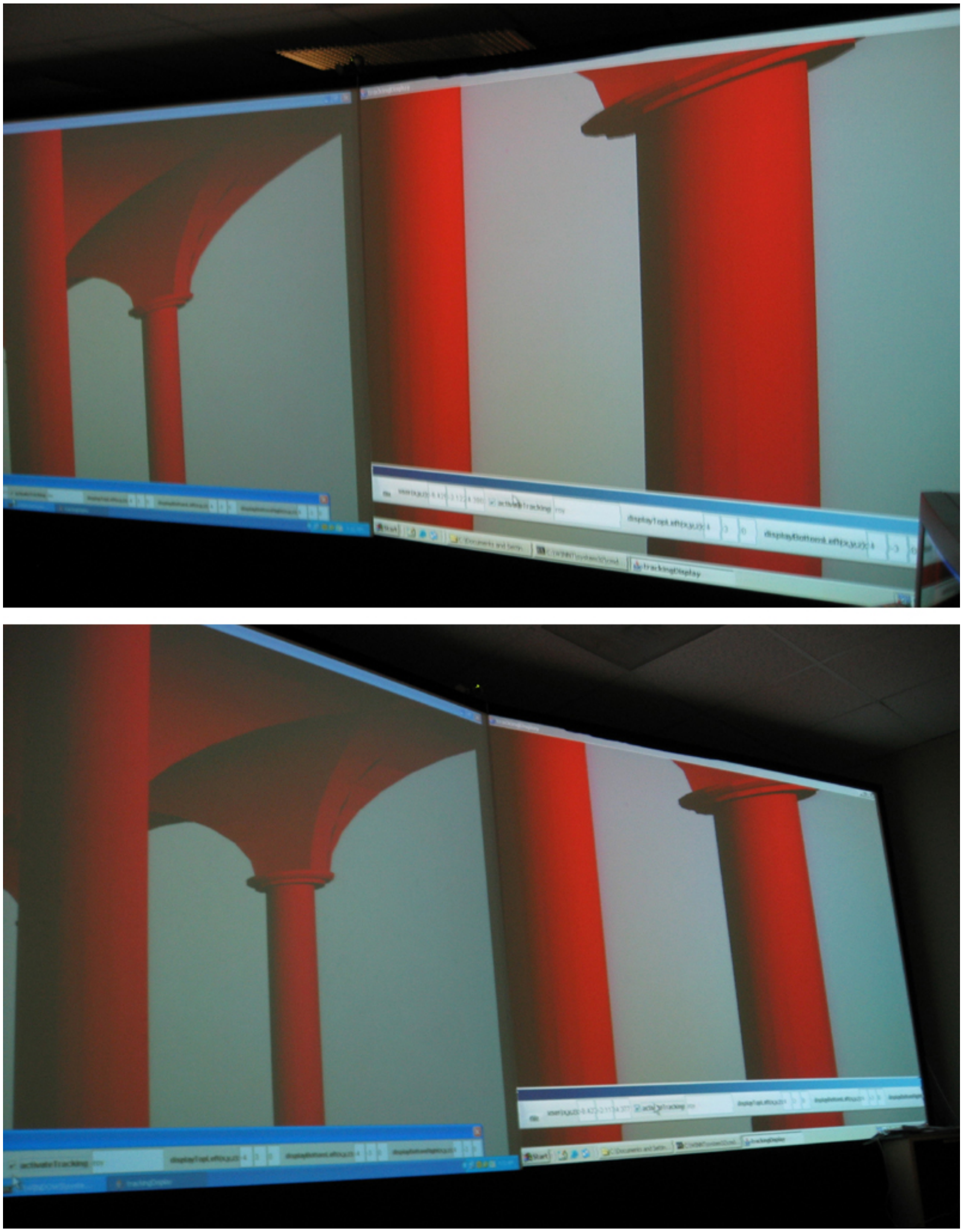


Figure 4.11: Two images of the same display compare the view of a wide angle perspective image from incorrect (top) and correct (bottom) locations. Note especially the marginal distortions that are evident in the column capitols in the off-axis view. These same columns do not appear warped from the correct viewing location.

Chapter 5

Conclusion

We have shown that perspective images can geometrically recreate what a viewer might see if they were looking at a real scene. This is because they can send the same light rays to the eye that a real scene would send. But this is only the case when the perspective image is viewed from the correct location. When viewed from an incorrect location, the image may appear distorted. The effects of scalar distortions that result from incorrect viewing location adversely affect a viewer's ability to judge space and scale in an image.

We documented the history of the development of the perspective image in order to illustrate that the artists and architects of the Renaissance were sensitive to viewer location. From this study, we renewed our emphasis on the relationship between viewer, image, and scene in order to reduce the distortions often evident in wide-angle perspective images.

We created a computer graphics application that presents a view-dependent perspective image. Our application uses a skew camera model to draw a unique perspective image that is based on the location of a viewer. In order to find the viewer location, we implemented a camera based tracking system. Multiple displays were tiled to create a wide-angle perspective image that does not suffer from distortions that are typically inherent in the display of wide-angle images.

We successfully created perspective images that do not suffer from the distortions that are commonly present in wide-angle perspective images. We did this by using a skew camera that uses viewer location data to present an image that geometrically recreates what the observer would see if they viewed a real scene.

Appendix A

Depth Perception

Here we provide a more comprehensive description of the different effects that combine to give the observer of a scene a sense of depth, distance and relative scale.

This is important to take into consideration during the design and implementation of computer graphics applications for architects. For example, the effects of aerial perspective would provide little indication of depth to a user looking at a scene in close proximity. Binocular disparity and convergence would provide important information for someone tying flies, but would be less important when viewing larger architectural spaces. Under normal circumstances, all of these cues agree with each other and lead to a coherent interpretation of the surrounding three-dimensional world.

A.1 Interposition

Commonly known as occlusion, this cue describes that if one surface blocks the view of another, it is seen as nearer. This is one of the most powerful depth cues, and is seldom overridden by any of the others. This depth cue is accomplished in computer graphics viewing by using the z-buffer for occlusion culling.

A.2 Relative Size

Objects of similar size at different distances cast retinal images of different sizes. This was first formalized when Euclid observed that the same object subtends a different visual angle depending on its distance from the observer.

A.3 Height in the Visual Field

Objects that are higher in the visual field are perceived to be farther from the observer than are lower objects.

A.4 Motion Parallax

When moving, the relative distance of objects determines the amount and direction of their movement in the retinal image. The most common example of this phenomenon is the perception of a landscape while travelling in a car or train. Objects that are closer move more rapidly through the retinal image of the moving observer than do objects that are more distant.

A.5 Binocular Disparity

A typical person's eyes are 6 centimeters apart. As a result, what we see through one eye is slightly different than what we see through the other. We have learned to correlate the two retinal images of an object or scene and to thereby judge depth. This cue is most effective at close distances.

A.6 Accommodation and Convergence

If an observer focuses on an approaching object, his eyes will axially converge. Information about the level of convergence is relayed to the brain from the eye muscles. For this reason, the usefulness of convergence cues drops off with distance until it is virtually useless at distances of greater than ten meters. This is because differences in the axial rotations of the eye diminish with distance.

A.7 Size Constancy

Also includes shape, brightness and color constancy. If an observer knows the size of two objects, his familiarity will override relative size cues. These are fundamentally learned cues and break down when the observer is unfamiliar with the actual size of an object.

A.8 Aerial Perspective

The passage of light through air causes effects that give indications of relative depth at extreme distances. The composition of the atmosphere effects the level to which this becomes relevant. Haze and a hue shift are common indications that aerial perspective cues are in effect. For example, the popular line ”Purple mountain's majesty” refers to the aerial perspective cue that the mountains, because they are seen at a great distance appear shifted in hue.

Appendix B

User Location Tracking

Here we discuss in more detail several different methods that are currently used to provide user location data. We do this so that the reader will understand the concepts, limitations and current capabilities of each approach.

B.1 Mechanical Tracking

Mechanical tracking relies on a direct physical link between the target and an environmental reference point. This usually consists of two or more rigid mechanical pieces connected using electromechanical transducers that measure the movement of each link of the arm. Although not inherently limited, range is typically one cubic meter. As range increases, the mechanism can become quite elaborate. For this reason, mechanical tracking is used primarily for specialized applications and is considered too cumbersome for widespread application in graphics

B.2 Inertial Tracking

The use of inertial tracking in computer graphics was delayed because the system relies on gyroscopes to determine orientation. The technology was initially used to track the orientation of ships, airplanes and submarines. But until the advent of *microelectronic mechanical systems* (MEMS), which can place gyroscopes and accelerometers on a microchip, it's application to graphics was limited. Three orthogonally positioned gyroscopes attached to the target position measure the changes in orientation. Three accelerometers (with orientation similar to the gyroscopes) measure the acceleration vector in body space. This is then converted using the current rotation matrix as determined by the gyroscopes. The gyroscopes indicate the direction you are facing, and the accelerometers tell you how fast you are accelerating in that direction. The position relative to a known starting position can be determined using this information.

Inertial systems have no line of sight limitations and low latency [1]. They are not affected by interfering electromagnetic fields or ambient noise. Measured velocity and acceleration allows for prediction of up to 50 ms into the future. The major problem with the approach is drift. Bias error by the gyroscopes of 1 milliradian can cause the reported position to vary from the actual position at a rate of 0.0098 m/s2 (this is 4.5 meters of position error over 30 seconds). Even very good gyroscopes will drift by this amount over time. One solution has been to periodically re-calibrate the gyroscopes' location in a hybrid system[2].

---

[1] Less than 2 ms

[2] An acoustic system is often used to reset the gyroscopes location because its slow update rates, which alone are insufficient for most applications, satisfy the needs of a hybrid system.

B.3 Magnetic Tracking

Thee orthogonally oriented magnetic sensors provide a 3d vector describing the unit's orientation. It is possible to use the naturally occurring magnetic field of the earth, but many graphics application require more accuracy than this method provides. For this reason, standard practice is to actively induce excitation using a source unit. Magnetic tracking is commonly used in immersive graphics because it provides six degrees of freedom and the user-worn component can be small. It is subject to distortion if there are conductive materials in the environment that would affect the shape of a magnetic field. There still exists the need to wait for an excitation to subside as well as time for a measurement of the ambient field, but measurements are typically made at a rate of 100 per second or more. Because the magnetic field has an inverse cubic falloff, range and resolution fall off rapidly. Practical range for most commercial systems is less than three meters. Because magnetic fields pass through the human body there are no line-of-sight limitations.

B.4 Acoustic Tracking

In acoustic tracking, time of flight measurements calculate the distance from the emitter of an ultrasonic pulse to a receiver. By using several emmitter/reciever pairs, one can find three-degree-of-freedom location information. This approach is limited by multi-path reflections because the receiver cannot distinguish between direct and indirect waves. Because sound travels relatively slowly, it is necessary to wait until the first pulse arrives, measure it, and then wait for the pulse's reflections to dissipate before sending the next pulse. This process takes five to ten milliseconds depending on the environment. Because of this limitation, update rates are bound to below ten hertz. A higher frequency pulse increases resolution and reduces the effects of environmental interferences, but reduces range. Acoustic trackers are also sensitive to humidity, wind and ambient noise.

B.5 Optical Tracking

A general characteristic of optical tracking is that light from a source is measured by a receiver. Often filters are used that are calibrated to the source to restrict the measured light to the invisible range (infrared). There are two basic divisions of optical tracking solutions: analog and digital. In the first, analog photo-sensors measure the quantity of light reaching a receiver from an emitter. Position data can be calculated if several emitter/sensor pairs are used and ratios compared. Alternatively, analog position sensing detectors can give the position of the centroid of a light.

The second approach is to use a CCD measuring device such as that found in a standard digital camera. A standard implementation of this type includes two or more cameras calibrated to view a tracking area. Each camera looks for a predetermined target. When the cameras locate the target, they can supply a vector that describes the direction from the camera to the target. Two or more vectors from different cameras are used to calculate the target location. More cameras can be used to improve accuracy and to compensate for occlusion.

B.6 Ultra Wide-Band Tracking

Ultra wide-band (UWB) tracking uses time difference of arrival measurements to provide location data. Active tags attached to the target emit energy over the entire spectrum from DC to several GHz. This allows the emissions to look like low-level background noise and reduces the effects of multi path reflections. Most UWB tracking systems are scalable and have the potential to precisely track thousands of targets over millions of square feet of indoor space. The most common current application is used for tracking assets within hospitals and industrial facilities [FRB03].

Appendix C

Digital Perspective Images for Architectural Design

We introduce some of the concepts that were necessary to understand in order to address some of the problems with designers use of computer generated perspective images. We approached the problem from several background areas: architectural design, computer graphics, perceptual psychology and history.

C.1 Design

The work of an architectural designer relies on various modes of representation. Architects often use drawings, sketches and models to convey information about a space that they envision. They also use similar aides to give themselves an idea of what the space might be like as part of their own design process.

Design is inherently an iterative process. In designing a space, an architect creates representational imagery that is then critically examined. Based on what is learned from this exercise, the imagery is adjusted, new images are created, and the process repeats itself. Each iteration brings an increase in the specificity with which the designed space is represented.[Pic98] The perspective image has a long history as a tool in the design process. This is because it can give the designer a sense of what the space might look like from a particular vantage point. All of this is true whether the imagery is created using a computer or by some other method.[Dur02]

Two things that would improve the design process are the quantity and quality of its iterations. The computer has greatly increased the speed with which perspective images can be created. It follows that this should therefore increase the *quantity* of possible iterations in the design process. Designers claim that the computer generated image does less in the area of improving the *quality* of each iteration.[1][BK03][Law99] By improving the ability of the architect to see space and scale in a perspective image, we hope to improve the quality of each iteration of the design process.

The perspective image has long been a powerful tool used by architects to convey a sense of what a built space might look like. This is because such an image can geometrically recreate what a viewer might see if they were looking at the real scene. Historically, the perspective image has been used to extend the space in architecture by giving the impression of a three-dimensional scene on a two-dimensional surface. This is evident in the efforts during the Renaissance to extend the perception of space through the painting of perspective frescos. The power of a perspective image lends itself readily to use by architects, both as part

---

[1] To be fair, the greatest increase in the quality of each iteration that is possible in the design process could be improved most by improving the designer. That issue at first seems to be outside the scope of this thesis. We hope to improve the designer's abilities by presenting a better image.

of the design process as well as a tool for selling ideas to clients.

Since the development of computer graphics, the computer has been used to draw perspective images of scenes. An early impetus for the development of these techniques was the potential for their use in the field of architectural design. With the advent of computer graphics technology came also an outcry from the design community claiming: *we can't use this for design.*[Row99][Tur01] And so the computer graphics image has to large extent been relegated to the "selling side" of the design process. A three-dimensional space is conceived and reworked by the designer using traditional media (Figure C.1). It is usually after this initial creative period that the use of digital technology is introduced. And even at that point, its use is largely restricted to the creation of marketing images and construction documents. While there are instances where digital perspective images are used as part of the design process, their potential still far outpaces the utility of their actual application. This may be for various reasons. One claim is that the current paradigm in graphics technology provides an inadequate sense of scale within a space. This is partly due to the fact that the field of view is narrow and the viewer is stationary in front of an image (Figure C.2). The increasing size and quality of display devices make this paradigm less realistic.[McL02] An increase in display size provides for the use of wide-angle perspective images. Unless the viewer's location is known, these wide-angle images present even more distortions in the perception of scale than do those presented on smaller displays. Our goal is to create computer generated perspective images that can be more effective as an interactive part of the design process.

Figure C.1: The first year design studio at Cornell University is housed in Rand Hall. Here a student works on a design problem and is surrounded by drawings and models. These aide in understanding and conveying an idea of what the space and scale of the built work might be like.

Figure C.2: Adjacent to the design studios at Cornell's Rand Hall is a student computer lab. Students, professors and practitioners shy away from using a digital environment in design for a variety of reasons. Foremost among them is that the interface between designer and work is lacking - as is evident in this image.

C.2 The Perspective Transform

The process of creating a perspective image varies depending on the discipline. As a result, architects and those who study computer graphics understand perspective differently. Architects usually construct a perspective image by choosing up to two vanishing points toward which lines that are parallel in the scene and recede in depth from the observer converge in the image(Figure C.3).[2] This is the common method used to teach about perspective in architecture schools.[Tol00] This understanding of the construction of a perspective drawing takes into account the placement of the viewer within a plan drawing of a space.

Because it is tedious to construct a complex perspective image with any degree of accuracy or control, the computer is often used to aide in their creation. A computer graphics perspective image is created by transforming a three-dimensional scene which is then orthographically projected onto a two dimensional display. The process of distorting the scene geometry so that it can be used to create a perspective image is known as the *perspective transformation*. An important concept to keep in mind is that perspective projection results in a geometry that is modified, but still three-dimensional(Figure C.4). While both methods result in the same image, the architect frequently fails to understand the power of three-dimensional perspective geometry.

---

[2] One point and three point perspective images contain lines that converge to one or three vanishing points. These types of drawings are used less often by architects than is the two-point perspective.

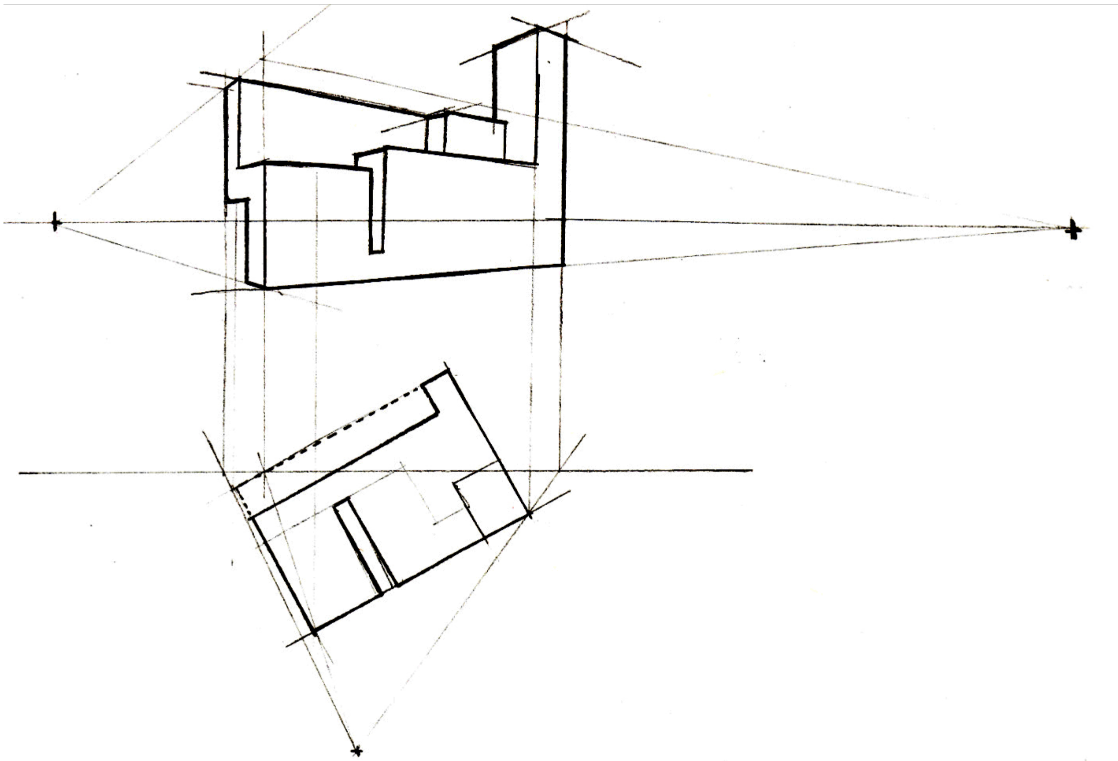


Figure C.3: An architect's understanding of perspective is usually based on the method that is used to construct an image. A plan view (bottom) with accompanying station point and image plane defines the locations of vertical lines in the image(top). Horizontal orthogonal lines converge toward one of two vanishing points(left and right of the image).

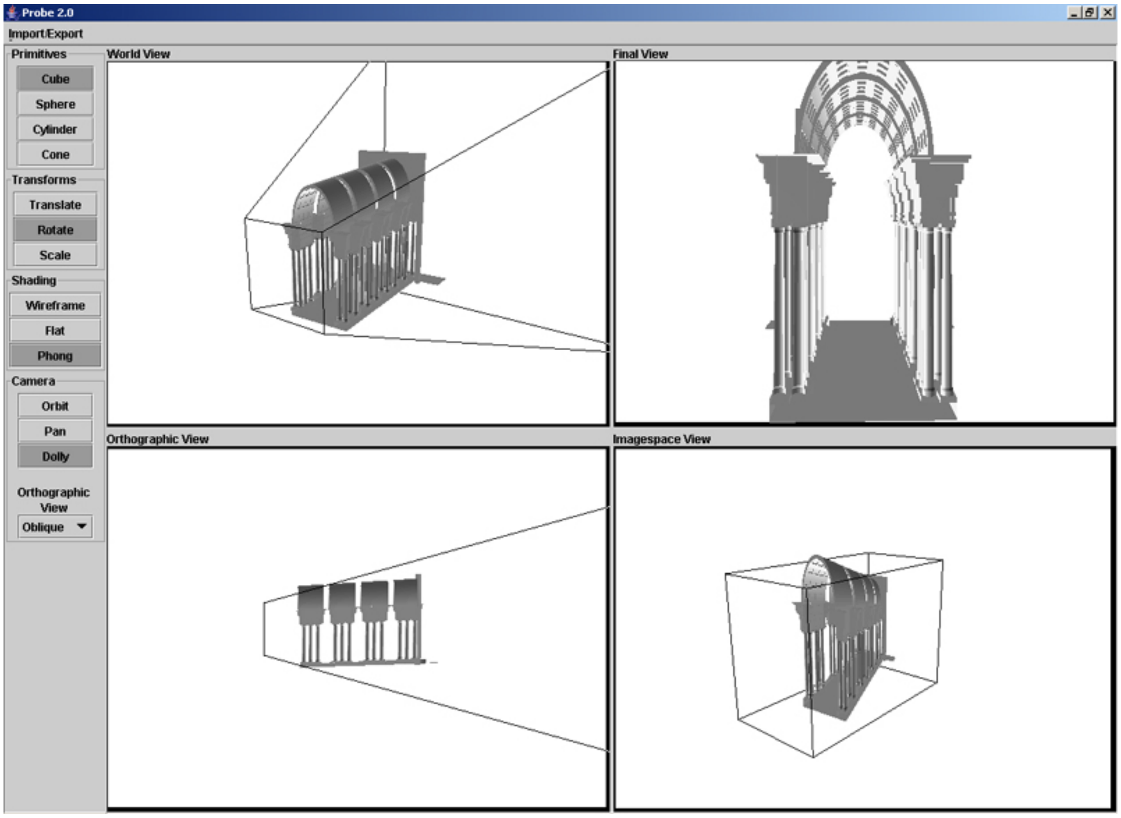


Figure C.4: Probe is an application developed to teach the concepts of perspective transformation in computer graphics. All four views display the same scene: The two images at the left show the scene before the application of a perspective transformation. Note especially that the camera's frustum in these two views is shaped like a truncated pyramid. The lower right image shows the scene geometry after the perspective transformation. This geometry is used to arrive at the final image (upper right).

Bibliography

[BK03] Julio Bermudez and Kevin Klinger. Digital technology and architecture – white paper. *the Association for Computer Aided Design in Architecture*, 2003.

[Bor04] Ann Marie Borys. Lume di lume: A theory of light and its effects. *Journal of Architectural Education*, 57(4):3–9, May 2004.

[Bro99] Frederick P. Brooks. What's real about virtual reality? *IEEE Comput. Graph. Appl.*, 19(6):16–27, 1999.

[CKKP] Jaeyong Chung, Namgyu Kim, Gerard Jounghyun Kim, and Chan-Mo Park. Postrack: A low cost real-time motion tracking system for vr application.

[Dur02] Fredo Durand. An invitation to discuss computer depiction. In *NPAR '02: Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering*, pages 111–124, New York, NY, USA, 2002. ACM Press.

[FR80] James Farber and Richard Rosinski. Compensation for viewing point in the perception of pictured space. In *The Perception of Pictures*, volume 1, chapter 4. Academic Press Inc., 1980.

[FRB03] Robert J. Fontana, Edward Richley, and Joann Barney. Commercialization of an ultra wideband precision asset location system. *IEEE Conference on Ultra Wideband Systems and Technologies*, Nov 2003.

[Kem90] Martin Kemp. *The Science of Art.* Yale University Press, 1990.

[Kub86] Michael Kubovy. *The Psychology of Perspective and Renaissance Art.* Cambridge University Press, 1986.

[Law99] Bryan Lawson. Fake and real creativity using computer aided design: some lessons from herman hertzberger. In *C&C '99: Proceedings of the 3rd conference on Creativity & cognition*, pages 174–179, New York, NY, USA, 1999. ACM Press.

[Man70] Antonio Manetti. *The Life of Brunelleschi.* Pennsylvania State University Press, 1970.

[McL02] Charles W. McLaughlin. Progress in projection and large-area displays. In *Proceedings of the IEEE*, volume 90, pages 521–532, April 2002.

[OM04] Takehiko Ohno and Naoki Mukawa. A free-head, simple calibration, gaze tracking system that enables gaze-based interaction. In *ETRA'2004: Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications*, pages 115–122, New York, NY, USA, 2004. ACM Press.

[Pic98] Moreno Piccolotto. *Sketchpad+ Architectural Modeling Through Perspective Sketching on a Pen-Based Display.* Cornell University Thesis, 1998.

[Ren02] Ronald A. Rensink. Internal vs. external information in visual perception. In *SMARTGRAPH '02: Proceedings of the 2nd international symposium on Smart graphics*, pages 63–70, New York, NY, USA, 2002. ACM Press.

[Row99] Colin Rowe. Architectural education: Usa. In *As I Was Saying*, volume 3. MIT Press, 1999.

[RWC$^+$98] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: a unified approach to image-based modeling and spatially immersive displays. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 179–188, New York, NY, USA, 1998. ACM Press.

[Sal01] Tomas Garcia Salgado. Anamorphic perspective and illusory architecture. In *Programme of Generative Arts Symposium*, Milan, May 2001.

[SCCZ86] Leslie Stroebel, John Compton, Ira Current, and Richard Zakia. *Photographic Materials and Processes.* Focal Press, 1986.

[Shi02] Peter Shirley. *Fundamentals of Computer Graphics.* A K Peters, 2002.

[TDM01] Osama Tolba, Julie Dorsey, and Leonard McMillan. A projective drawing system. In *SI3D '01: Proceedings of the 2001 symposium on Interactive 3D graphics*, pages 25–34. ACM Press, 2001.

[TL99] K. Talmi and J. Liu. Eye and gaze tracking for visually controlled interactive stereoscopic displays. In *Signal Processing: Image Communication*, volume 14, pages 799–810, 1999.

[Tol00] Corey Toler. *A Computer-Based Approach for Teaching Architectural Drawing*. Cornell University Thesis, 2000.

[Tor77] Harry Torczyner. *Magritte: Ideas and Images*. Abrams, 1977.

[Tre99] Camillo Trevisan. The borromini gallery in palazzo spada, rome: Ideal regular model and deformed real model abacus of the perspective deformations. In *Prospettiva e Prospettive*, may 1999.

[Tur01] Ziga Turk. The reasons for the reality gap in caad. In *Proceedings of the 19th Conference on Education in Computer Aided Architectural Design in Europe*, pages 156–160, 2001.

[WF02] Greg Welch and Eric Foxlin. Motion tracking: No silver bullet, but a respectable arsenal. *IEEE Comput. Graph. Appl.*, 22(6):24–38, 2002.

[YGH$^+$01] Ruigang Yang, David Gotz, Justin Hensley, Herman Towles, and Michael S. Brown. Pixelflex: a reconfigurable multi-projector display system. In *VIS '01: Proceedings of the conference on Visualization '01*, pages 167–174, Washington, DC, USA, 2001. IEEE Computer Society.

[YM04] Jungyi Yu and Leonard McMillan. A framework for multiperspective rendering. In *Eurographics Symposium on Rendering*, 2004.

[ZB95] Denis Zorin and Alan H. Barr. Correction of geometric perceptual distortions in pictures. In *SIGGRAPH '95: Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 257–264. ACM Press, 1995.