Perspective in Two Dimensions for Computer Graphics

by

Elodie Fourquet

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2012

© Elodie Fourquet 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Computer graphics perspective is based on photography, the pin-hole camera model. This thesis examines the perspective as practiced by artists, who develop the picture geometry within the planar surface of the canvas. Their approach is flexible, depth is simulated with planar composition as the primary geometry. Renaissance artists discovered construction methods to draw the foreshortening of realistic pictures: the construction of a tiled floor in perspective was fundamental.

This thesis presents the framework, a computer program, I developed to create the perspective of pictures based on the geometry practices of artists. Construction lines on the image plane simulate the 3D geometry of the pictorial space; cartoons of foreground elements are manipulated in 2D within the picture perspective; projected shadows, examples of double projection, are also included. A formalism, reformulating algebraically the straight-edge and compass evaluations, generalizes the planar geometry that solves the challenge of depicting 3D. A revised *Painter's algorithm* produces the occlusions between the picture elements from sequencing them from their definitions on the canvas.

Acknowledgements

I would like to thank all the people that made this extraordinary journey and its finale possible.

I am grateful to my supervisor Professor Stephen Mann who took me over as a Ph.D. student on short notice and with no compromise. Steve, I am thankful for your guidance, trust and patience without them this outcome, of which I am glad and proud, would not have happened. Your corrections have made this work better.

To all my committee members I am thankful to have made my defence a memorable day. Your appreciations and insights from a broad spectrum are the greatest honour one wish to terminate graduate life. Professors Richard Mann and George Labahn, I am in debt for your support during my years of resilient effort and your careful corrections on this final document. Professors Rick Haldenby and Pat Hanrahan, I am privileged for your comments and opinions on this dissertation work and on its perspectives for the future.

This dissertation of my dreams would simply not exist if I had not meet Bill Cowan. I am very appreciative that Bill commented and corrected this manuscript in various stages of disarray, before copy editing its latest versions using up at least a full bottle of ink! From Bill I have not only learned how to pursue creative research, but also he has showed me how to live. I am most grateful to Bill for teaching me how to persist under all the unknown circumstances of life with thought, integrity and kindness. Bill, thanks for the cappuccinos and espressos, for the books and operas, for the finest food and French wines, and for Celeste and Florence among many other great things.

There are further people I like to thank for making graduate life joyful, interesting and easier.

First, I am lucky to have meet a special friend, Alma Juarez Dominguez, to share ups and downs of graduate life torments. Alma, I want to express my deepest thanks for your moral and technical support. You have been a precious help for me since our first all nighters completing projects, to those last moments of defending a dissertation. You are my little fairy.

Furthermore, I would like to thank all the CGL members for making a place to work and to eat ice-cream and coffee hours, so stimulating and pleasant. Eoghan, Jeff, Alex, Cherry, Eugene and Phil, thanks for sharing with me your knowledge, your fun and our weirdness. Miro, Glenn, Jimmy and Berkant thanks for those moments of fresh air soccer and stress relievers drinking. Thank you to all my students for your engagements during lectures and outside them, they have been essential breaks from the isolating life of research.

I would also like to thank my family, my dad and mother, my sister and little brother, who believed and supported me in this crazy pursuit. Merci maman de me parler de si longues heures au telephone et de m'écrire si souvent: ton esprit est plein d'histoires qui donnent à penser et me font sourire chaque jour.

Finally, I am grateful to the Natural Sciences and Engineering Research Council of Canada, the Ontario Graduate Scholarship in Science and Technology, and the Cheriton School of Computer Science at the University of Waterloo, all of which financially supported the research presented in this dissertation. Furthermore, I would like to acknowledge the Library of the University of Waterloo and the Tri-University Group on which I relied to obtain many books, some obscure from the past, others most recently published. Thank you for your great collection and for letting me have hundred books out over many years!

Contents

List of Figures

Chapter 1

Introduction

This thesis revisits the process of creating pictures in perspective using a computer. Drawing pictures in perspective has a 600 year history. Computer graphics is largely based on a narrow and recent fragment of that history: the era of photography. As a result, computer graphics fails to take advantage of the deep insights into picture making realised in the works and writings of the great realist painters. From the Renaissance beginning, painters understood that a flexible approach to geometry, simulation of 3D within a 2D composition, provides the rich mixture of metaphor and reality needed to create a picture that lasts for centuries.

Renaissance artists invented perspective and understood its relationship to geometrical optics. As early as 1490, Leonardo da Vinci made a sharp distinction between *artificial perspective*, 'the painter's projection of forms on to a plane' [65, p. 49], and *natural perspective*, which describes geometrical optics in the actual world. Indeed, 'perspectiva artificialis' (also known as 'prospettiva accidentale') is 'the Renaissance Latin term for the artist's linear perspective' [32, p. 198], not to be confused with 'perspectiva naturalis', which refers to the mathematical laws governing natural vision [37, p. 21]. The linear perspective used in computer graphics is rigid and mathematical, an implementation of natural perspective wholly distinct from the artificial perspective of the artist.

This thesis examines artificial perspective, as realized in the practice of artists. Following Renaissance practice, I use construction lines on the image plane to simulate the 3D geometry of pictorial space, and add cartoons of foreground picture elements manipulated in 2D within the picture perspective. Projected shadows are also constructed as examples of double projection. I often contrast geometric optics with my work, but natural perspective is not used in my pictures. My pictures only simulate 3D on the canvas, deduced from the intersection of construction lines fully on the image plane.

(a) (b)

Figure 1.1: A pair of pictures that illustrates the difference between natural perspective and artificial perspective. (b) is *The Geographer* (c. 1668) by Vermeer. In (a) the globe has been modified to obey geometrical optics [42].

1.1 Motivation

Artists privilege 'making over matching' when they create, as their goal is to interest, intrigue, and arouse the 'beholder' [47]. Computer graphics, in contrast, prizes scientific precision. This section illustrates the difference, using a painting by Vermeer. The dominance of science in computer-generated pictures is often criticized; seeking an alternate approach, this thesis considers the practice of artists to balance the scientific influence on computer graphics.

1.1.1 One example

Consider the two images in Figure 1.1. Which is more convincing? Which looks most right? Which communicates better?

Figure 1.1 (b) is a painting by Vermeer, created with artificial perspective. Figure 1.1 (a) is a copy, with the globe atop the closet modified to accord with geometrical optics, the natural perspective of computer graphics [42]. The oblong shape is the image of an off-center sphere on the image plane of a pinhole camera. However, we desire to view the globe as portrayed in (a), a simple sphere. Paradoxically the mathematical 'correct' perspective looks 'weird'.

Why did the artist reject natural perspective? Because artificial perspective freed him from replicating optics when doing so looks wrong. Vermeer's painting narrative is a geographer analyzing maps.

1. The globe is important to the narrative context. It associates the compass, held by the geographer in his study, with the wide world.
2. Globes are spherical.
3. Ergo, the depicted globe should look spherical.

A spherical shape is more easily recognized as a circle than as the elongated ellipsoid of physical optics. Artificial perspective gave Vermeer the flexibility to paint the best representation of a globe, *qua* globe.

1.1.2 Criticisms of Computer Images

Artists overruling mathematical optics in favour of perception is not rare. This characteristic in their practice (Chapter 3) is a challenge to computer graphics.

Computer-generated images have often been criticized for their overall look. Some notice a mechanical look and a lack of visual harmony [27]; others see a 'high-tech' look, especially art historians [35]. Currently, non-photorealistic rendering seeks to alleviate such criticisms, by simulating artistic media and styles [48].

But imitating artists and their media is insufficient. Mills noticed that 'realism', as defined by physics, has little relevance to 'realism', as defined by artists, that 'there will always be fundamental differences between looking at images and looking at the real thing' [79]. Mills refers to the work of artists as dissimilar as Constable, Rembrandt and Monet, to show that the artist's goal of evoking thoughts and feelings is qualitatively different than the scientist's goal of reproducing a projection of reality[1]. Picture-making is a mode of expression. A picture by its nature is not a substitute for something real, but a thing in itself.

[1]There is quite a difference between a portrait and a mug shot.

Mills recommends paying attention to art historians, such as Gombrich [47], and to perceptual psychologists, such as Hochberg [86], for improving education and practice in computer graphics. In particular, he emphasized Gombrich's idea of 'the beholder's share', the idea that successful depiction requires an effort shared between the artist, who creates the image, and the beholder, who completes it in his or her imagination. This idea is essential for understanding the shortcomings of computer graphics images, which are so overdetermined as to leave little to the imagination.

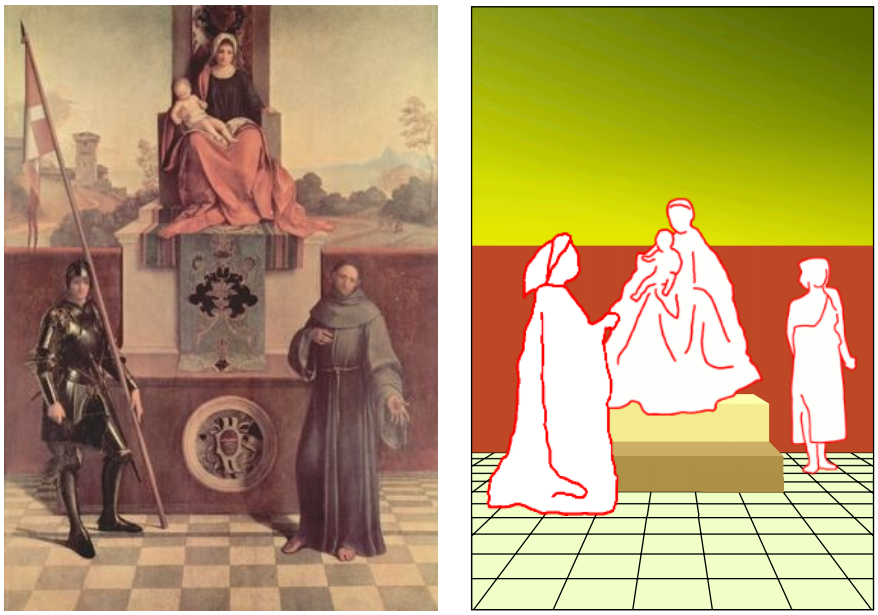
1.1.3 Artistic Practices

Computer graphics research followed on the path indicated by Mills. For example, a critical difference was noted: in traditional media, artists learn how to turn each limitation to advantage [20, 30]; graphics, in contrast, erases limitations with layers of technology.

Returning to Vermeer's painting, the following comments open the discussion of Renaissance practice which is continued in Chapter 3. Vermeer's globe has a local central perspective, and is probably an independent study, in the style of Renaissance cartoons. But many elements of the painting do not have independent perspectives: the furniture has a geometric global coherence. Lines of the closet, of the window, its pane dividers, its frame and sill converge to a single point: the principal vanishing point, located just out of the painting frame on the right. In addition, there is another type of projection, subtly present in the painting: projected shadows, which are not part of shading, but a picture element projected onto a surface by occluding the light from the window. For example, on the globe, a shadow cast by its stand is visible; the geographer's thumb casts a shadow on the book; and on the window sill, the pane dividers seem not straight owing to the reflected shadow.

The geometric structure responsible for coordinating these elements into a satisfying composition is hidden. *The Geographer* is from the late seventeenth century, long after the geometric innovations of the early Italian Renaissance had been subsumed into artistic practice. Those innovations were centred on the projective geometry of perspective. In practice, they appear in paintings as a tiled floor, a prominent horizon and/or cartoons. Figure 1.2 (a), the *Castelfranco Altarpiece* (c. 1505) by Giorgione, shows the explicit geometry as part of the picture.

To understand better how this geometry supported composition, I built a framework in which a user can compose in the style of the early Renaissance, creating first the view geometry then populating the space with two-dimensional cartoons, producing a true 'collage of different constructions' [36, p. 131]. Figure 1.2 (b) shows an exploration of the *Castelfranco Altarpiece* with the two perspectives [36, p. 234–236] collapsed into one, and the planar arrangement of Saints



(a) (b)

Figure 1.2: (a) *Castelfranco Altarpiece* (c. 1505) by Giorgione. It contains a structural component important in the development of perspective: a tiled floor. (b) is a schematic representation created with my framework: the tiled floor is fundamental to the picture perspective; the foreground figures, each of which is depicted in its local perspective, are inserted in the perspective picture; the tiled floor help positioning, sizing and determining occlusions.

Liberale and Francis broken. My framework, while built on the explicit geometry of the early Renaissance, applies to later realist paintings. Another work of Vermeer, *The Astronomer* (1668), portrays another arrangement of much the same foreground elements in a similar room to *The Geographer*, as if the painting geometry had been planned using my framework.

1.2 Renaissance Space

Computer graphics software provides two distinctive methods for creating images, which are examined in Section 1.3. Both contrast with the methods of Renaissance painters, who accepted the flatness of the canvas, which is necessary for composition. But Renaissance painters also invented the geometric methods of representing picture depth, as described below.

1.2.1 The Flat Canvas Surface

This thesis restricts itself to one element of picture-making: a painting's geometry[2]. Renaissance artists brought to life the flatness of the image surface and the rigidity of the static viewpoint. In doing so, they created a geometry, now called projective geometry, able to simulate 3D realistically on a 2D surface. But they understood that the flat picture is processed by human vision, which responds to two dimensional features.

Mathematically, linear perspective is a many-to-one mapping from 3D to 2D. Thus, a 2D picture cannot correspond to a unique 3D world. Furthermore, the picture need not correspond to any physically possible 3D world for it to be visually intelligible. Artificial perspective guides the realistic representation of 3D objects, without excluding the freedom necessary to create a convincing composition supporting the picture's narrative.

The Euclidean geometry of antiquity was important to Renaissance artists, who studied it, mastered it and used it in their compositions, often as regular 2D geometric figures (Section 3.2). Imitating artistic practice by accepting the flatness of the picture plane and turning its limitations to advantage is necessary for computer graphics to achieve the richness of expression available to realistic painting.

1.2.2 Mathematical Innovation

Renaissance perspective pictures depend on important advances in geometry. Beyond Euclidean geometry is the geometry of vision, how things *are seen* rather than how they are, which was understood and brought to fruition in early Renaissance artistic practice. To do so, Euclid's parallel axiom, stated below in an adapted modern version (in the style of [99, p .22]), had to be generalized.

Modern parallel axiom. *For any line $\mathcal{L}$ and any point P not on $\mathcal{L}$, there is exactly one line through P that does not intersect $\mathcal{L}$.*

In effect, the novel constructions of Renaissance paintings are based on the observation that on a picture, parallel lines cannot be distinguished as lines that never intersect. Contrary to 'reality' parallel lines in perspective do meet: the appearance of the tiled floor in perspective is a demonstration. This observation is a breakthrough, '[v]isually obvious but mathematically mysterious' [99, p .94]. In fact, what defines parallel lines is where they intersect, at infinity. The horizon is a line at infinity, where parallel lines in a realistic picture must converge.

[2]In the terminology of Willats, I am focusing on one of the two elements of representation: *the drawing system*, as opposed to *the denotation system* [107].

Euclid's parallel axiom was indeed challenged by the geometric innovation required by realistic paintings. Specifically, to construct the tiled floor Renaissance artists had to recognize the following facts, which are true in any view of the plane [99, p. 94].

1. Straight lines remain straight.
2. Intersections remain intersections.
3. Parallel lines remain parallel or meet at the horizon.

The invariances lead to practical rules, such as any point at the intersection of two lines remaining at the intersection. They permit the construction of realistic pictures, developed on a plane without compromising 2D composition. The 3D world does not dominate, its representation emerges from the few rules to which the canvas lines conform.

Long before projective geometry was formalized in the nineteenth century, the perspective practices of Renaissance artists demonstrated by example the properties of the projective plane and its relation to the 3D world represented in a realistic picture. In effect, the line intersection constructions can be reformulated as projective geometry computations in the real projective plane, $\mathbb{RP}^2$, as shown in Chapter 6.

1.3 Graphics Software

To situate my research in a familiar context, I here discuss the treatment of 3D and 2D spaces in current graphics tools. Neither modelling software, in which a complete 3D scene is defined by the user, nor imaging software, in which 3D appearance receives no explicit support, easily replicate the 2D geometry developed in the Renaissance to create and organize realistic images.

1.3.1 3D software

One class of tools for making pictures, especially sequences of pictures for animation, is *modelling software*, exemplified by Houdini™, Blender™ and Maya™. It supports a user who creates 2D output from fully-modelled 3D content.

The construction of 3D models is difficult, expensive and time-consuming. Thus, they are generally created for mass production, or when many views are needed. Also, they must be articulated, capable of taking many poses. To make up a scene several 3D models are sized, posed and arranged in 3D space. Only then are the camera and lights positioned.

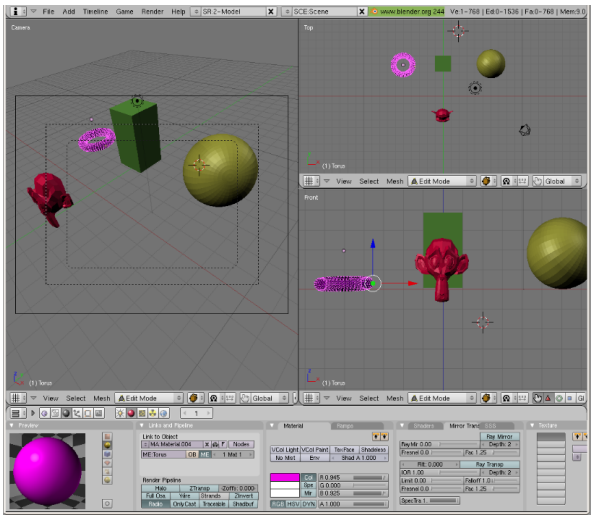


Figure 1.3: Graphical user interface of a 3D modelling software, Blender. The user works in the 3D virtual space through a 2D interaction model. Using many views of the scene he makes local adjustments, often decomposed into 2D motions along world axes. On the left is a camera view of the scene; on the right the top view and frontal one.

Thus, modelling environments require the maker of a single image to specify far too much. Furthermore, they lack the high level communication using natural language with which artists control their models.

More is less. 3D modelling software requires the user to set and tune hundreds of free parameters in creating a 3D scene. While doing so they work between 2D and 3D: the interface (Figure 1.3) and the image are 2D, but the definitions and manipulations are all 3D. The ongoing dominance of 2D in graphical user interfaces for desktop interaction demonstrates that 3D has yet to be integrated into the cognitive capabilities of the general user. The problem is unlikely to be inadequate designs: mental rotation experiments suggest that 3D objects present a high cognitive load.

The modelling burden is not lightened by view transformations in the rendering pipeline. Pinhole projection[3] is applied rigidly, subjecting every 3D vertex to the same transformation, which does not always have predictable effects. Thus, if the composition is unsatisfactory, the artist must choose among modifying 3D object definitions, 3D object arrangement, lighting, and camera placement. Many iterations are usually necessary to create a good image.

[3]More complex systems, mimicking lens aperture or depth of field for examples, are available. But are as they are rigidly applied, control of distortion and composition are no easier with them.

The iterative process is similar to the practice of artistic photography. The photographer chooses objects and lights, then arranges them with respect to the camera position to create a well-composed image. However, the photographer has a significant advantage, rendering by eye. Through the viewfinder he sees continuous changes in the image as objects, lights and camera move through space. The result is a much denser sampling of possible images than can be provided by 3D software, making possible more targeted refinement of the image. The greater malleability of virtual objects in 3D software is a drawback rather than an advantage. Modelling software is more demanding than photography because the physicality and interactions of objects are virtual.

Traditional 3D design example. An example from life that shows the need for high level controls and for doing only what is necessary appears in a documentary film by Sydney Pollack about Frank Gehry (*Sketches of Frank Gehry*, 2006), who is rightly proud of his computerized architectural practice.

With his assistants, he constructs by hand a maquette of the new building using easily handled materials such as sheet metal, cardboard, balsa wood and so on. Only when the design is complete in a miniature physical form, is the maquette hand-digitized by a technician using 3D software. The 3D mesh then is used to specify detail and create assembly plans, articulating the building in terms of structural elements and textured surfaces. Thus, it is no surprise that contemporary artists, who are highly trained in traditional media, commonly reject 3D modelling software, preferring the relative predictability of 2D software.

1.3.2 2D software

2D computer graphics is the basis of a variety of 2D tools, grouped together as *imaging software* and exemplified by Microsoft Paint™, Adobe Photoshop™ and Adobe Illustrator™. Imaging software supports a user working directly on the 2D canvas. Thus, the user can focus on 2D composition, but gets no direct support for simulating 3D.

Flexible traditionally-based context. The 2D software digital environment, which resembles non-digital media, leverages drawing and painting skills that have been practised from infancy, simple, well-tested methods descended from five centuries of artistic tradition. On the surface of a virtual canvas the virtual equivalents of physical media, pencils, pens, brushes and paints, are intuitively manipulated using conventional interaction models. Direct interaction with the final product makes it easy to control composition.

2D software seems complex only owing to the quantity and specificity of tools. Using only a few tools, users obtain useful results quickly, learning as they work. The combination of early rewards with steady progress is both satisfying and motivating for the user.

In addition, the digital medium supplies extensions that are well-defined and cognitively immediate. Affordances such as undo, repeat, and transform, speed up manual actions and forgive errors. Thus, computation provides advantages that attract artists to 2D software as a complement to physical media: the computer playing the role of the many assistants found in Renaissance workshops, in decreasing the tedium of repetitious manual tasks.

Limited depth support. However, unlike 3D software, 2D software only poorly supports the geometry of realism. 2D transformations, which traditional artists perform manually are automated, but the absence of depth reference requires the user to measure. For example, 2D transformations that warp or distort need to coincide with the image perspective. A user who desires 3D-like foreshortening has the choice between tedious calculation and interminable adjustment by eye.

Thus for a picture created with 2D software to have the appearance of depth requires many specific user actions. The software provides derisory assistance for depth arrangement and no guide for perspective. A single feature that provides minimally for depth is the layer paradigm, which provides a partial depth order, perceived depth being supplied by occlusion.

A counter-example. 2D software is designed for creating images, which are dominantly 2D representations[4]. But it is also used for realistic images, including complex 3D representations.

Kevin Hulsey is a professional illustrator who uses Adobe Photoshop™ and Illustrator™ to create cutaway technical illustrations. In Figure 1.4, which depicts the volume of the boat in 3D, he created an elaborate perspective structure, drawing hundreds of construction lines on the flat canvas in order to present the boat in an attractive 3/4 view, with strong apparent depth and 2D composition. Although a computer was used, the geometry of the boat was done by hand, using methods closely resembling those of Renaissance artists: each construction line was drawn individually, followed by local depiction of the boat's elements.

Hulsey's example indicates that there are substantial gains in working on the image plane when creating an image requiring easily-perceived depth. The effect would be too mechanical if it were created by 3D software working from plans. The desire for 2D control explains the preference of digital artists for imaging software when creating realistic images from scratch.

[4] 2D representation is defined in Section 3.1.1.

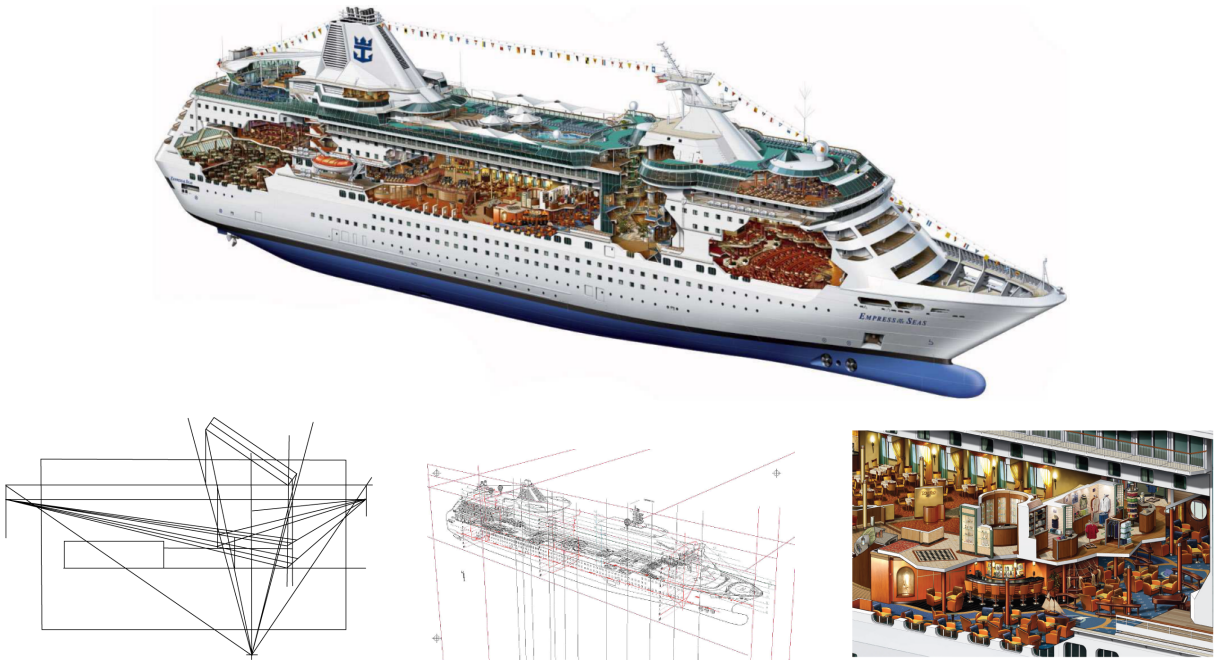


Figure 1.4: Kevin Hulsey's illustration, *Empress of the Seas' Cruise Ship Cutaway* (Software: Adobe Creative Suite™; Illustration Time: 960 hours) [60]. The lower images show the initial construction lines, setup for the 3/4 view, a later stage of the geometric construction, and a cutaway showing interior rooms.

Ideally, 2D software would incorporate tools to support pictures like Hulsey's. An example might be automated construction lines, by which artists control the geometry of depth. Since Adobe Creative Suite 5 (released May 2010), Illustrator has perspective drawing tools, including construction lines to define the geometry of vector drawings. Simple perspective constructions leverage a technical user's knowledge and skills while helping a novice to acquire the rudiments of perspective. The application I developed for this thesis has been designed in this style, providing support for the appearance of depth on a 2D canvas.

Computer tools for creating images are thus divided between those that are fundamentally three dimensional and those that are two dimensional. With 2D software, planar composition guides the depiction from start to finish, with 3D software, it is 3D veracity. This thesis opposes this division, aiming to fuse working on the image plane with ease of simulating depth. By examining the practice of artists important depiction concepts can be identified. Considering them tools that serve both the expression of 2D and the construction of 3D might be designed.

1.4 Overview

This thesis describes the framework I built based on Renaissance painters' practice: 2D geometric algorithms mimic their constructions in the digital medium of the computer. The goal is to ease the combination of simulation of depth with composition on the canvas, which Renaissance paintings show to be possible. It is a prototype that permits the user to investigate interactively the spatial relationships in a picture when the perspective is artificial.

1.4.1 Framework

Renaissance constructive methods were developed to satisfy the demands of 2D composition and 3D arrangement. On the canvas, artists evaluated the positions, sizes and appearances of objects to fit realistically within a scene that portrays depth. Their methods produce spatial realism directly on the flat canvas where 2D composition is intuitive. Renaissance painters simulated 3D with a perspective that is neither simple nor unitary, but plural.

My framework implements and generalizes geometric constructions used by artists creating artificial perspective. Specifically, artificial perspective is based on two types of elements, both of which are present in Figure 1.2 (b): one type simulates the architectural volumes constituting the scene perspective, the other populates the foreground space with object representations that are largely independent of the scene perspective.

In my framework, canvas construction lines merge two sets of landmarks, the salient points of the composition and the guidelines of the perspective: the former with a planar pattern; the latter with a depth structure on the ground. The *tiled floor*, i.e., a chequerboard pavement, which is visible in Figure 1.2 (a), embodies the principal perspective to which the geometric volumes adhere to. The framework's user controls the picture perspective through the geometry of the tiled floor. Foreground objects are created in local central projection, which avoids the undesired distortions. Projected shadows can also be geometrically constructed by the framework under the direction of the user specifying the canvas location of the illumination.

These Renaissance practices conform to human perception. The viewer's freedom to move when observing a picture is taken into account, which requires flexibility in perspective, flexibility that allows the 2D composition to be primary.

Developing the framework enables to address in this thesis the following questions.

- Can the practices of Renaissance artists inspire new ways of developing the geometry of realistic depth?

- Are objects that are fully defined in 3D essential for realism?
- Is it possible to work in 2D, simulating 3D in the manner of realistic painters?
- Can image plane composition be made easier without compromising the viewer's comprehension of 3D arrangement?
- Is it possible to provide more effective 3D-like interactions using 2D manipulations?

1.4.2 Scope

This thesis proposes a new approach to creating realistic pictures. Being an initial step in a new concept of image creation, the results should not to be compared to state-of-the-art computer graphics images. When imagining realistic computer graphics through the perceptions of the great realist painters, geometry comes first. Achieving a geometric framework rich enough to encompass their practice is the required base. Simplifications are made to focus on the essence of the constructions discovered empirically within traditional art practice. The purpose is to understand better the intrinsic spatial properties of depiction. Other problems, such as lighting, colours, surface details, brush strokes and so on, are also important, but they need a geometric foundation and are therefore beyond the scope of this thesis.

For painters, the working space is the image plane. Thus, in this thesis it is the domain of all computations: 2D computations evaluate the 3D simulated scene. All the points used to calculate the perspective picture are on the 2D canvas; this restriction is central to my proposal. The geometry of scene objects is represented at least by a planar contour at most by the volume occupied, to the exclusion of fine detail. Thus, realism, as used in this thesis, is narrowly defined; *schematic representations*, one of which is shown in Figure 1.2 (b), better describe my intent and results. Schematic representations absorb most intelligibly the spatial attributes of a picture: the 3D arrangement of the objects depicted and their 2D composition.

Moreover, as in early Renaissance paintings, the relation of view and scene, as configured on the canvas, is treated rigidly. Most schematic representations are organized by central perspective: the principal architectural volume of the scene is presented with its front face dominant, which is suitable for axis-aligned scenes. Early Renaissance painters favoured such scenes and views, producing symmetrical geometry in the picture.

In developing schematic representations I bring to computer graphics the geometric realism of painters. At present realism in computer graphics is based on photography, and qualified as 'photo-realism'. Simulating physical optics produces a particular signature; schematic representations have their own signature.

This limited scope enables me to 'respect both the way that artists work and the limitations of software' [75]. Using it I can more easily learn and formalize the geometric construction used by artists to compute canvas locations in practice. Simple and highly structured arrangements allow us to isolate the geometrical invariants that exist when 3D is flattened into 2D. Artists learn them intuitively through long hours of experimentation. Like all humans they have a sense of order in space that is immediate, intuitive and correct. Understanding it better is a source of inspiration for computer graphics.

1.4.3 Contributions

The computer framework developed in this thesis help us to understand, test and explore the geometry of realist paintings by creating schematic representations. The framework differs from previous software by being based on Renaissance artificial perspective, directly constructed from evaluations on the canvas surface. Achieving the framework relies on several key contributions.

The panel abstraction. Panels individually separate foreground objects, each of which is drawn using an independent local perspective. The panel geometry, chiefly the attachment point, easily integrates the content within the perspective space of the picture.

The tiled floor. The tiled floor is the depth structure used to construct and arrange the perspective picture. Among the tiled floor Renaissance methods, the essence of its perspective construction is identified so that the canvas straight lines relate succinctly and elegantly to the 3D geometry in play.

Artists' construction lines. Geometric shapes are constructed in perspective by algorithms that evaluate intersections among canvas straight lines, as artists have done for centuries.

Algebraic formalism. Projective geometry on the plane is used to reformulate the construction algorithms algebraically: each sequence of line intersections is represented by a 2D matrix. An analytic formalisation of the Renaissance constructions is possible using Cartesian coordinates: 2D homogeneous coordinate evaluations can simulate the geometry of realistic pictures, which make the approach simple to integrate into 2D software.

Projected shadows. Projected shadows, which provide spatial cues, are also generated by constructive geometry and formalized in 2D algebraic geometry.

2D Painter's algorithm. Given the geometric context, a rendering based on a partial order among the tiles is introduced. The algorithm enumerates all valid sequences that produce correct occlusion.

Within these contributions identifying the substance of artists' tacit knowledge is a recurrent challenge. It must be made explicit definition if it is to be encoded algorithmically.

1.5 Organization

The remainder of this thesis is organized as follows.

Chapter 2 describes computer graphics research related to this thesis. Image-based research similarly mediates between the interplay of 2D and 3D, most often through photography. Fine art has also inspired a variety of computer graphics research: styles and 3D reconstruction being two important goals.

Chapter 3 reviews the geometry of Renaissance painting. It presents arguments that substantiate the importance of composition and the nature of geometric constructions executed on the canvas. Chapter 4 then presents the Renaissance construction line methods on which my framework is based. The tiled floor in perspective is described in detail along with a few examples of more elaborate constructions, including some of projected shadows.

Chapter 5 presents my framework. The components of a schematic representation are described. The geometric algorithms that construct and manipulate them using construction line intersections on the image plane are given. The components and the algorithms are based on the geometry and the canvas constructions of realistic artists (Chapters 3 and 4). Then, Chapter 6 presents an algebraic formulation based on 2D projective geometry to compute the evaluations made in Chapter 5 with construction lines: the formalism derives the 3×3 matrices for the depiction task. Chapter 7 then explains the painter's algorithm I devised for my framework to render schematic representations.

Chapter 8 discusses schematic representations achieved with my framework, many inspired by paintings of the Renaissance. Chapter 9 presents the contributions of this thesis and avenues for future research.

Chapter 2

Related Work

This thesis investigates two important themes of computer graphics research: mediation between 3D scene and 2D image, and inspiration from the practice of artists. Thus, this chapter presents the general themes and research ideas of computer graphics as they relate to this thesis.

2.1 Mediation Between 2D and 3D

A realistic picture mediates between three and two-dimensional spaces. Traditionally artists process the dimensional transfer as they lay down each brush stroke on the canvas, each of which has a correspondence in the world artists see or imagine. The computer enabled new aspects of the mediation between 3D and 2D when creating a realistic picture. The virtual canvas permits 3D to exist behind the image plane and as the user interacts with graphics software. However as described in Chapter 1, the availability of 3D throughout the process of creating an image is a debatable gain at least for non-expert users.

For this reason, research in computer graphics has developed many ways of mediating more effectively between 2D and 3D. For example, a variety of 2D input can be interpreted as determining 3D or at least the appearance of it. Simplifying the interpretation of user input is an important idea in sketch-based modelling and image-based rendering systems. The challenge is to avoid limiting the user's expressivity by taking into account their cognition. This thesis borrows ideas from sketched-based approaches and more importantly from image-based research. But first, it is worth mentioning that even true 3D software has concepts emphasized in this thesis.

2.1.1 Improvements in 3D Software

Computer-aided design (CAD) systems for example have been shown to require investment on the part of the user: perceptual challenges such as 3D blindness [87] or weak spatial intuition [63] need time to overcome. 3D software is generally complex, 'so different from the traditional method', that it is reserved to people 'simply trained to use it the way it is written' [75]. Such criticism has motivated the exploration that is reviewed in the next section. However, efforts have been made to facilitate learning of 3D software by integrating cues of human perception and methodologies of drawing practice.

For example, to help users manipulating 3D objects scattered in the virtual space of CAD, a multi-scale grid on the ground and an augmented connection to the ground for each object, have been investigated [46]. In the scope of this thesis, it is most interesting that simple visualization features making use of the dominant plane of 3D perception, the ground, have improved user performance.

Furthermore, incorporating other concepts close to artists' practices have been recently considered in CAD systems. Schmidt et al. addressed input complexity by designing an interface for analytical drawing [95], based on engineering technical design. In their work, related concepts, such as vanishing points, guide the construction of a 3D scaffold. The scaffold is the constraint structure for inferring the 3D location of subsequent curved input that creates a 3D volume. This work is a natural extension of sketch-based approaches, when integrated into 3D modelling packages. Thus, it is sensible to next describe the ideas behind sketching systems together with image-based ones, both of which serve in making or simulating 3D easier.

2.1.2 2D Abstraction

Much research in computer graphics concerns itself with reducing the dimension of user input without reducing the precision of output. This objective has a long history: Sutherland's first sketching system [101] has had numerous successors.

Sketching Systems. The sketching abstraction is important to this thesis: an intuitive paradigm based on drawing practice is universal. Sketching systems imitate physical strokes on paper to receive controls: 2D gesture grammars specify 3D. While the interaction is intuitive, extrapolation and interpretation are necessary for the process to generate 3D from 2D input. Thus, sketch-based approaches trade the accuracy and completeness of 3D modelling software for ease of use.

In comparison, this thesis takes the pencil and paper paradigm more literally. Narrowing the results to pictures in perspective, the geometry follows the actual ruler and compass constructions artists have developed through centuries, which determines the capabilities offered by the framework. Another characteristic of the framework is that all algorithms must be evaluated completely with 2D. Because the result is a single view, neither extrapolating nor interpolating is required for the algorithm computations. A further abstraction however is present in this thesis, because it follows artists' practice. Foreground objects are handled as the cartoons of Renaissance painters, a strong abstraction that is perceptually powerful compared to physical realism (see Chapter 3).

The second parallel between sketch-based approaches and this thesis is the distinction between two types of form. Until recently[1], sketching systems focused either on rigid objects fully defined by 3D or on malleable objects whose primitives are of organic origin. Sketchpad and its successors, from SKETCH [110] to the freeware SketchUp belong to the former, while inflating mechanisms like the one introduced in Teddy [61], are specifically designed for volumes whose shapes are free-formed. In my approach, the rigid geometry of architecture is handled differently from the natural, organic—usually human—forms found in the foreground.

Unlike sketch-based systems, the purpose of which is to create a 3D model, my framework supports creation of 2D pictures only. There is an advantage in this restriction. An image corresponds in general to an infinite number of 3D scenes, or to no possible 3D scene[2] Using layers and anchors in 2.5D, Rivers et al. showed it is possible to go beyond 3D using a combination of few flat representations [94]. This work showed there exists an alternative to working 3D without compromising expressive potential.

Image-Based Rendering. The second type of systems that use 2D input usually with 3D algorithms is image-based rendering. Modifying the 3D appearance of photographs, controlled by simple user manipulations, has gained tremendous interest, since digital photography ubiquity. Image-based rendering exploits the interplay between 2D and 3D spaces. Its usual goal is to alter 2D content, often photographic in origin, while photo-realism is sufficiently maintained for the modification to be unnoticed.

Combining the strengths of computer graphics and computer vision, many capabilities have recently been achieved. In a photograph, displacing or modifying objects now encompasses

[1]Only lately, combinations of 2D silhouettes with computational solid geometry has surmounted this strong division between the types of form [93]: three fixed 2D views, each a silhouette, are used to model complex shapes.

[2]In *The Representation of Things and People* Julian Hochberg contrasts the general replacement for most depicted object and the phenomenon in inconsistent pictures [86].

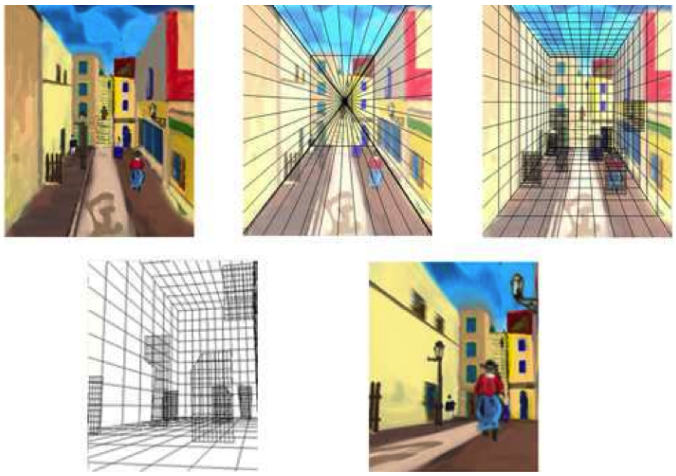


Figure 2.1: Process flow of *Tour Into the Picture* [58]. Upper row: input image; spidery mesh; foreground objects on billboards. Lower row: camera positioning; output image.

a flexible range of outcomes [112]. In addition, researchers have inserted non-photographic effects, such as pop-ups [56], extensions beyond the frame [97], and cutouts [92] and walk throughs, into photographs most often to compensate for limited perception of depth.

The common attribute of image-based rendering is partial reconstruction of the geometry of the 3D scene, which is needed to create modifications that satisfyingly blend into the unchanged content. The degree, the method, the precision of geometrical reconstruction vary from one system to another. At one extreme photo-realistic rendering is emphasized: the camera geometry is inferred, and all visible surfaces are given 3D locations. At the other the goal is minimizing computation while maintaining salient perceptual features: good-enough abstractions are the key to improvement. Techniques for preserving perceptual salience are based on ideas that are important to this thesis.

The abstraction provided by billboard representations is the simplest approach. Billboards separate the foreground scene objects, which are clipped away from the background. In four instances, billboards have been used within a coarse 3D scene geometry, which is significant to my research.

First, Horry et al. showed how to decompose a photograph into the scene perspective volume and the foreground objects, albeit with the assistance of the user [58]. A 'spidery mesh' filling the volume in the photograph identifies the geometry of background planes aligned with the principal vanishing point. Billboards are used for foreground objects, as illustrated rightmost of the upper images of Figure 2.1. Once the simple reconstruction is complete, some

changes of view and restricted animations of the camera are possible, showing that the photograph has been separated into a coherent set of elements, as illustrated lower row of Figure 2.1. The simple 3D representation derived from a photograph is closely related to my approach. The main difference is that in Horry et al. the 3D simulation is derived, i.e., the mesh structure and the billboards are pasted onto an existing photograph. My framework, in contrast, assists the construction of a 3D-like space in which elements are placed to create an image composition.

'Harold: a world made of drawings' [16] demonstrates billboards in a 3D virtual space. A three dimensional terrain, with a hemispherical backdrop, is filled with billboards for objects and billboard bridges for connected objects, so as to allow the scene to be navigated without a full 3D definition of the world. Similarly, the commercial software products of Toon Boom [105], an enhanced version of Flash, supports the creation of 2D animation by positioning elements in parallel planes at various depths so that a camera can move in depth within the scene.

Finally, 'out of bounds' images, which are compelling, depth-rich images generated from photographs [97], are based on concepts relevant to this thesis. Two strong depth cues, occlusion and shadows, are exploited; foreground objects are separated from the background by placing them on billboards. Object masks, created with user assistance, enable foreground figures to cast shadows onto planar surfaces using a 'roughly' calibrated geometry. A major feature of image-based research, which differentiates it from my thesis, is the centrality of computer vision. Improved camera calibration for better image understanding and more extensive statistical training on image types, which is important for computer vision to move forward, is necessary for image-based rendering. Interestingly, in parallel with my thesis, the assumption of downward gravity and the separation of the foreground elements from the principal perspective surfaces is leading to significant advances in image understanding [57].

These few examples demonstrate that a simple representation provided by billboards can approximate depictions of 3D objects provided that 3D navigation is sufficiently constrained. In my application, which supports creating perspective pictures, representing foreground elements as billboards is ideal. Arranging billboards in a simulated 3D space goes beyond their previous uses, because it relaxes the close coupling between a billboard and the view, providing simultaneous interaction with image plane composition and depth arrangement. Billboard rotations are possible, up to limits that depend on object geometry [40]. Thus, billboards, which I formalize as panels, are an integral part of my schematic representation.

While all those image-based approaches, which mediate between 3D and 2D information to change a picture, are inspiring, two fundamental differences with this thesis approach are worth emphasizing: the context and the purpose are different.

In the photographic alteration context, compositing mechanisms, such as are needed to isolate an object prior to moving it, are essential: graph-cut mechanisms help content extraction and seamless pasting into the modified photograph [62]. Because my framework composes each picture starting from a blank slate, image processing prowesses is unnecessary. 3D geometrical truth is less important than composition.

The purpose of my thesis is to incorporate artists practice into geometric evaluations to simulate the 3D scene of a picture. Therefore, my approach replaces the camera reconstitution of computer vision by the practices painters have developed to create realistic representations. Among image-based systems, the image warping approach of Carroll et al. control is closest to this thesis: construction lines, are used to modify a photograph, supporting more explicit perspective than is realisable with a camera [12]. Both construction lines and vanishing points act as constraints guiding the optimization that transforms the perspective of the photograph.

To summarize, my work while neither a sketching system—as 3D is only simulated—nor an image-based one—as input image are not modified—relates to and borrows from these two fields of research: 2D abstraction is inherent to them and to this thesis.

2.2 Analogy with Shadows

The previous section described reducing the dimension for user input while maintaining the realism of 3D. An analogy can be made with projected shadows. Their appearance is a mediation between 3D and 2D: depending on illumination characteristics in relation to blocking volumes, all of which placed in 3D space; volumes cast shadows onto planes, all of which are transferred to the 2D canvas. There are unanswered questions about how to control predictably the double projection of rendered shadows [66]. This thesis, by including projected shadows, participates in improving illumination controls and simplifying shadow appearance, following on ideas the research described below.

As was previously done for scene geometry, I new describe attempts to simplify or enhance the control of shadows. In 3D modelling environments direct control of highlights and shadows instead of the light that creates them improves the search for illumination parameters that suit the user [88]. Furthermore, the visual function of shadows depends on the context in which they occur. For scientific realism precise optical physics is necessary. But, in other contexts flexibility and practicability are most important. In cinematography, for example, visually rich lighting is essential. Thus, research in computer graphics has developed lighting controls

Figure 2.2: Images from *Stylized Shadows* by DeCoro et al. [24].

and features inspired by real-world techniques. Imitations of the tricks and cheats of lighting designers enhance illumination effects in computer animated films [9, 84]. This research path demonstrates that by-passing the rules of physical realism can be a benefit to visual expression. This thesis emphasizes the perceptual richness that lighting can provide.

Another source of inspiration for dealing with the complexity of shadows is observing their appearance in art. DeCoro et al. proposed a shadow stylization algorithm [24] based on the painting shown in Figure 2.2 (b): a large vase the projected shadow of which is lacking in geometric detail. The vase shadow emphasizes the separation of the foreground figure from the background ones without drawing attention to itself. Unlike the usual computer graphics shadow algorithms, which emphasize exactness, DeCoro et al. imitate artists' shadow representations by inflating the shadow matte of ordinary 3D algorithms to elide details. Their results resemble the projected shadows generated in my framework. The difference is that in my framework shadows are produced by the shape of the painted object, so starting from an approximation they do not need further simplification.

Reinforcing DeCoro et al., computer vision research into shadow compositing shows that projected shadows can be evaluated within the weak 3D geometry retrieved from a photograph [11]. Like the computer vision approach, my framework includes projected shadows within the perspective of the 2D picture. The difference is that in this thesis the simulation and control of projected shadows is based on artistic practice. The projected shadow derivations (§ A.4) contributed in this thesis are already features of an image-based system that relights input images [59].

2.3 Inspiration from Art

Research motivated by art is a substantial part of computer graphics: non-photorealistic graphics, for example, is devoted to effects first realized by artists. Algorithms expanding the range of visual appearance in computer-generated images are numerous. The denotation system, in particular, has been emphasized: brush strokes, media effects, painting styles, and color schemes have been imitated [48]. For example, filters based on the dominant stroke characteristics of different painting styles have been developed: the texture of a photograph can be transformed accordingly [54] or the rendering of a 3D scene [76].

Automation of artistic effects is a success for non-photorealistic research and has many practical applications. But observing art offers another possibility, which is to learn from the practice of artists so as to improve the practicality and control of graphics tools. Such scrutiny of art might transform into lasting improvements to the expressivity of graphics software users, for users who artistic control in addition to automation.

Because this thesis focuses on the geometrical aspects of creating pictures, this section reviews research investigating artists' richness in spatial expression. Much of it was inspired by geometry of existing paintings, observing results rather than practices.

2.3.1 Multiple Perspectives

Paintings often deviate from the strict linear perspective of the pinhole camera. Inspired by this observation researchers have explored in two directions, one taking into account the complexity of visual perception, the other imitating the complex projection models of modern art.

The distortion that pinhole geometry produces for off-center spherical objects differs from the undistorted representation used by artists (Figure 3.7). It has long motivated researchers in computer graphics. Donath explained the nature of projective images, contrasting the physical reality of the geometry with the perceptual qualities found in art [26]. She encouraged a compromise between the two representations, carefully representing spatial.

In the same spirit Zorin and Barr, noting that mathematical faithful projections are not always visually correct, proposed re-projecting photographs to compromise between distorting spheres and bending lines [113]. To rectify the distortions of linear perspective, multiple perspectives have been introduced to bring mathematical calculations closer to perceptual expectations. Franke et al. experimentally demonstrated an aesthetic preference for multi-perspectives found in art over the single-perspective produced by the pinhole camera model [42].

Other research has extended the camera model to achieve modern, non-realistic scene projections. For example, Agrawala et al.'s [2] found inspiration in the compositions of Paul Cézanne [72]. Singh was motivated by multiple simultaneous viewpoints of a single object in Picasso's work [98]. This result was extended to nonlinear projection using multiple linear perspective cameras to render a complex animated scene Coleman [18].

These investigations increased the number of cameras. The user must manipulate many cameras in the 3D virtual space. As a result the number of parameters that must be adjusted to produce the desired image is further increased, making its creation that much more difficult. Yu and McMillan provided a general framework for preserving spatial coherence in multi-perspective rendering [108]. Mashio et al. focused on reducing the search space for the user; getting a set of images from representative viewpoints one per object; synthesizing a non-perspective image with a smoothing algorithm [73].

Other research concerned itself with images composed of objects projected from different points of view, using 2D images as input. Collomosse and Hall imitated cubist paintings using multiple photographs [19]. The composition is constructed from a set of photographs: areas are selected within them according to the presence of salient features. Those areas are then geometrically distorted to be composed in a single image, which was passed through a painterly filter. Zelnik-Manor and Perona created algorithms that produced Hockney-like collages from a collection of photographs [111]. Most interestingly, Meyer developed a custom image editing system to create multi-perspective collages from a dozen input photographs mapped to a 3D model: his results have some resemblance to those of this thesis [77].

The work discussed above was generally motivated by geometry observed in paintings, but often failed to engage directly with the goals of the artist. In realist painting the artist provides an experience to the viewer, largely through the relative positions of salient objects on the canvas, the painting's composition. And, without any compromise of the composition, the objects themselves are painted in projections that both show the essence of each object and provide a consistent simulation of 3D. In creating his compositions Cézanne neither traced rays across the canvas nor reproduced measured angles, both of which are part of multi-perspective rendering. Similarly, Picasso did not require an automated blending tool when assembling images into collages.

My interest is to avoid venturing into the artistic complexity of abstract representation, but to improve the expressiveness of realistic composition, based on a creation environment that is as natural and familiar to the user as possible. While automatic computations, mostly 2D homographies, occur in my work, they explicitly follow known practices of Renaissance artists, simplifying tedious manual tasks usually delegated to assistants. My 2D projective matrices

respond to the structure of the tiled floor in physical space and in the perspective picture. In my framework the user leads the 2D computations.

2.3.2 Painting 3D Space Retrieval

Investigating the geometry of Renaissance paintings has attracted attention in computer vision. Renaissance painting provides inspiring input for testing 3D reconstruction techniques. Art historians have been faced with spatial irregularities present in works of art; computer vision researchers are now participating in the debate, bringing new analysis methods.

Criminisi et al. analysed Renaissance paintings [22] using new techniques Criminisi developed for computer vision. Criminisi's algorithms were demonstrated by interpreting the 3D space, architecture, foreground figures and view parameters, which presumably defined the content portrayed on the canvas.

More recently, computer-aided geometric analysis of paintings has been used to investigate art history. For example, *Las Meninas* (1656) by Diego Velázquez is a painting with many open questions about the spatial arrangement of its subjects: the royal couple is reflected in the back mirror, and omitted from the visible part of the painting. Stork and Furuichi constructed a 3D computer graphics model of the figures within the room—the group of women, the painter partially occluded behind his large canvas, and the man through the doorway at the far back—as well as the viewer's space outside the painting [100]. Using it, they elucidate different postulates about the viewing configuration that may have been used.

Similarly the geometry of scenic architectural views of Venice painted by vedutisti[3] artists such as Pannini or Canaletto and of ancient Rome portrayed in Piranesi's etchings has inspired modern analysis. For example, a reverse engineered mathematical projection of a Pannini painting has been proposed [96]. Using her architectural training Rapp chose to use the *restitution* technique of Johann Heinrich Lambert, to explain Piranesi's innovative style of perspective manipulation, which permits him to control the visual impact of the etching on the viewer [90]. The resulting geometric representations combine different viewing positions, using creativity in use of perspective and composition where it is most essential.

While my research examines the space portrayed in planar depictions, I also observe the 2D geometry of paintings, so that my technique and purpose are different from the above work. My research does not try to retrieve 3D information, which my framework does not need. I use constructive geometry in the style of Rapp: 2D projective relations create a simulation of 3D.

[3] *vedutisti* is Italian for views.

My purpose is not to discover the 3D precision of a projection, but to understand geometrical patterns in the image plane that provide the 3D spatial coherence required for a convincing realistic 3D arrangement. This interest permits me to by-pass the assumption of a single pinhole camera, which is integral to the geometrical analysis of realist paintings. Instead, I enthusiastically embrace the pluralist approach to perspective found in Renaissance painting.

2.3.3 Artistic Practices

As discussed in this chapter, there has been a recent interest in considering artistic concepts and practices in computer graphics, especially in terms of geometry: vanishing point and construction lines are now an intuitive mode of manipulation [95, 12]. More significantly, those practical artists' concepts are now an integral part of Adobe Illustrator™, where it provides considerable assistance for the simulation of depth in vector representations. Based on perspective practices, new features have been incorporated that help users to depict depth on the flat canvas. In addition, placing of human figures in local perspectives has been investigated in 3D computer graphics research.

Adobe 2D software, Photoshop™ or Illustrator™, has taken several approaches to support designing the appearance of 3D, within the canvas environment and also outside it[4]. A decade ago, Illustrator™ added features to produce the appearance of 3D in 2D artwork: extrusion and solids of revolution simulation give a volumetric appearance to a flat object, with effects such as shading and rotation enabled, using automatic gradients and distortions. 3D painting followed. Few years later, Photoshop™ also took interest in depth and included a vanishing point tool to control the distortion of a texture and to map further textures onto it.

Subsequently, Adobe Illustrator™, following on favoured plug-in extensions and on Freehand graphics software, provided a different approach to perspective. Tools that assist users who wish to simulate the geometry of depth in an image, were introduced. Indeed, the recent adoption[5] of construction lines and vanishing points reinforces the promise of the results of this thesis [38]. One-point, two-point or three-point perspective grids, based on perspective representations of a cube, guide an automatic transformation of planar objects moved around on the canvas. The perspective grid establishes the foreshortening parameters in three perpendicular plane directions.

[4] Adobe Photoshop™ Extended editions have included features of 3D modeller software: initially restricted to editing partial texture and lighting of a 3D object on a 3D layer. Current version of the extended edition includes many features of a 3D modeller.

[5] In Adobe CS5, released May 2010, Illustrator™ new feature list includes perspective drawing tools.

Tutorial videos and texts explain how to use this feature. Their profusion suggests that the tool has been taken with great enthusiasm by the 2D graphics community. The tool implementation, in my observation, is little known. There is also a disconnection between previous 3D appearance features and the perspective grid tool: "3D tools are independent of the Perspective Grid tools and 3D objects are treated like any other object in perspective" (from the *Creating 3D objects* section [1]). This thesis addresses these two issues: a mathematical formalism is presented; and illumination effects accord with the appearance of the scene geometry. Notably, among many methods for simulating 3D in Adobe Creative Suite, the latest is similar to ideas investigated in this thesis.

Finally, the practice of artists is now taken into account when removing undesired distortion in 3D software. Zavesky et al. implemented a practice long taught in figure drawing courses in fine arts [109]. Their system uses an individual perspective for each human figure in the foreground of an architectural scene containing a strong principal perspective. They recreated *The Tribute Money* (c. 1426) by Masaccio, to illustrate how their results differ from pinhole camera images. My work is based on similar concepts in to Renaissance perspective. The difference is that my framework uses 2D images to compose a picture with perspective in simulated 3D, without requiring 3D models.

Chapter 3

The Geometry of Renaissance Art

This thesis develops a new approach to the geometry of realistic computer-generated images, which is based on the practice of Renaissance artists. This chapter describes how Renaissance painters worked, stressing the radical opposition between their conception of spaces and the geometry of 3D computer graphics.

Suppose for a minute that you are a painter in the early Renaissance. You have a well-developed heritage of medieval paintings that are structures in two-dimensions, their elements held in place by symmetry and hierarchy. But everyday you are encountering something new: the art of antiquity, which has an unexpected humanity and dynamism; the constructive geometry of perspective, which gives people, gods and saints a space in which to display their individuality; the philosophy of humanism, which gives value to the feelings and thoughts of every human being.

Six centuries later, in a time that is less innovative and intellectually exciting, we continue to assimilate your insights into our lives and thoughts. In this thesis I concentrate on your conception of geometry, and how it enabled you to bring so much into order. Introducing it into the thoughts and practices of the twenty-first century I hope to make it part of the mentality of practitioners of computer graphics. Why? Because seeing is as important as calculating, and I hope that this thesis will help my peers to see.

Renaissance artists choose how to convey the illusory sense of depth. They do not always follow strictly the geometry of physical optics, which determines a rigid mapping between the 3D world and the content of the canvas space. In particular, this chapter demonstrates that during the Renaissance, the golden age of perspective, artists were guided only lightly by physical optics: their rediscovery and elaboration of perspective was more than a mathematical

theory. For example they created a 'pluralist' approach to evoking depth on the flat canvas, with picture elements presented in one of two ways. Architectural and landscape define and adhere to a global perspective, receiving a rigid geometrical treatment that fixes the 3D space of the scene. Foreground bodies, which often have irregular shapes, are given individual local perspectives, often inconsistent with the global perspective. This practice creates a realistic appearance despite its geometric 'errors'.

3.1 Challenges

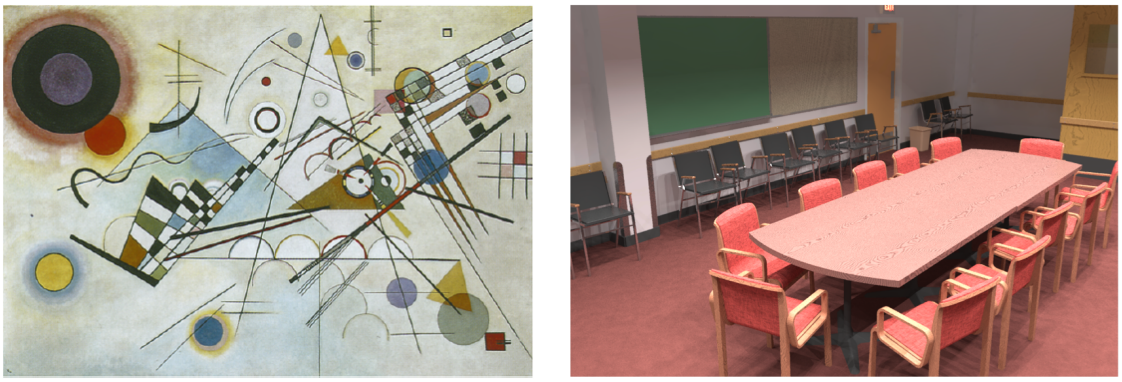
First, three spatial challenges inherent to the creation of realistic pictures are presented. Solving challenges like them guided Renaissance painters to their geometry practices. Such challenges still exist, likely confronting most modelling software users. Discussing them gives the context for the need of a new approach to perspective in computer graphics, inspired by the past.

3.1.1 Gluing together 2D and 3D

The first challenge for Renaissance painting was combining the spatial qualities of two and three dimensions into a coherent image. Every artistic representation of a 3D reality must be captured by the intrinsic planarity of the canvas surface. Creating images with harmony in two spaces, one 3D, the other 2D, is more difficult than making images that are exclusively two-dimensional or are purely three-dimensional, like sculpture.

Pure 2D representations rarely adhere to realistic geometry, such as most twentieth century and pre-renaissance art. One example is a painting of Kandinsky, shown in Figure 3.1 (a), where shapes on the canvas form a 2D geometric composition that mixes planar effects like visual weight, direction and balance [8]. In contrast, a representation that is solely naturalistic transcribes a realistic and consistent world geometry on the canvas, regardless of the planar consequences: identifying the geometric relationships among the depicted elements is foremost. However, such a representation, exclusively 3D, treating the geometry of the picture surface as incidental, often fails as art. The computer-generated image, shown in Figure 3.1 (b), which realistically represents a meeting room, is a defining instance of such 3D representations.

Indeed, when depth is not portrayed in a painting, it is rarely realistic; and when planar relationships are not present on the canvas, the picture is only naturalistic, and may be a random optical snapshot. For my purposes the distinguishing characteristic of Renaissance art is its embodiment of realistic depth together with 2D composition.



(a) (b)

Figure 3.1: (a) *Composition 8* (1923) by Kandinsky. (b) Computer-generated picture of a conference room (1997), realistic simulation by Ward [50].

In the realm of realistic painting, the canvas surface and the depth portrayed are two spaces that should blend together. Entanglement of depth within a flat structure is the *sin qua non* condition for spatial aesthetic in an image. Indeed, Hans Hofmann, the renowned twentieth century art teacher in two aphorisms insists on this critical aspect of picture creation [51, p. 65].

> Three dimensional objects in nature are optically recorded as two dimensional images. These images are identified with the two dimensional quality of the picture plane.
>
> The act of creation agitates the picture plane, but if the two dimensionality is lost the picture reveals holes and the result is not pictorial, but a naturalistic imitation of nature.

The first aphorism describes the optical relationship between objects in nature and the beholder's perception; the objects may be 3D, but the beholder's view is 2D. The 2D view is identified by the beholder with the 2D picture plane on which the artist works. The second aphorism describes the act of creation. The artist places items on the picture plane, possibly to suggest a 3D scene, but which primarily form a 2D picture. The 2D composition of this picture must be complete in the sense that it is visually satisfying. When holes, areas of the canvas where the composition is broken, appear in the composition, the beholder loses the 2D picture and sees only a naturalistic representation of a 3D scene.

In his comments, Hofmann is differentiating two types of picture. One is the particular 3D representation mentioned above, Figure 3.1 (b), a naturalistic representation of a scene,

where the beholder's goal is to perceive the 3D scene. In a random scene, virtually all possible viewpoints produce an image with items scattered indiscriminately on the 2D surface, thereby leaving 'holes'. In the second aphorism, the artist created picture is no more or less than a plastic alteration of the initially empty 2D picture plane. Effortful cognitive actions provide a composition that achieves plenitude within the boundaries of the canvas: the shapes and positions of portrayed objects define interrelationships that satisfy the eye's desire to find pattern when it perceives an object that is obviously cultural in origin. The picture itself is primary; the beholder provides the cultural reference that extends its meaning into the external world.

Renaissance paintings are pictures of the second type, both natural in the appearance of depth and well-composed in the pictorial space. Spatial attributes are portrayed harmoniously, in depth and in the picture plane. It is essential to identify the fundamentals in play by which geometrical relationships are created in and between the spaces.

Computer graphics research has not addressed the problem of reconciling realism with picture composition. Composition in a few instances has been considered to improve a realist rendering. The most notable works rely on photographic style guides, such as the rules of third or fifth, the golden ratio or canonical viewpoints to optimize the composition of snapshot. Gooch et al. use such heuristics to make more artistic the composition of a random image of a 3D object [49]; Lui et al. applied similar heuristics to crop a photograph [71]. Preferences in composition is a field where there is much to find out, as perception research indicates [70].

3.1.2 Undesired Distortion

A realistic image is a static representation from a fixed viewpoint. As a scientific problem, a realistic image aims to be a surrogate of viewing (with one eye) a 3D scene, but to artists the material substitution is peripheral to their intention. For artists, realism encompasses the complex cognitive responses evoked when experiencing a world recreated on a flat canvas.

Artists, faced with the flat and static picture plane, saw the limitation of a fixed viewpoint, which creates a geometric challenge as important scene elements away from the centre of the painting look unnatural because of the linear projection. In effect, rigid application of perspective results in perceptual distortion that depends on the visual angle [44]: wide angle photographs look 'weird', as shown in Figure 3.2. The distortion is particularly noticeable in closed curved objects. Painters, unlike photographers, can avoid optically correct projection when they sense its impropriety.

Renaissance painters noticed the perceptual distortion created by the mathematics of perspective projection. Both Piero della Francesca and Leonardo da Vinci studied the visual effect

Figure 3.2: A wide-angle photograph of a room (approximately 100° field of view) [113].

of perspectively projected columns [65]. Identical columns the same distance from the image plane do not project identically: their projected widths increase laterally, the further they are from the line of sight, as shown in Figure 3.3, which presents the geometry from a top view. Columns known to have the same radius do not appear the same. Painters felt free to draw them the same size on the canvas: regularity trumped realism. In such a scene, the mathematically correct image looks awkward and incorrect.

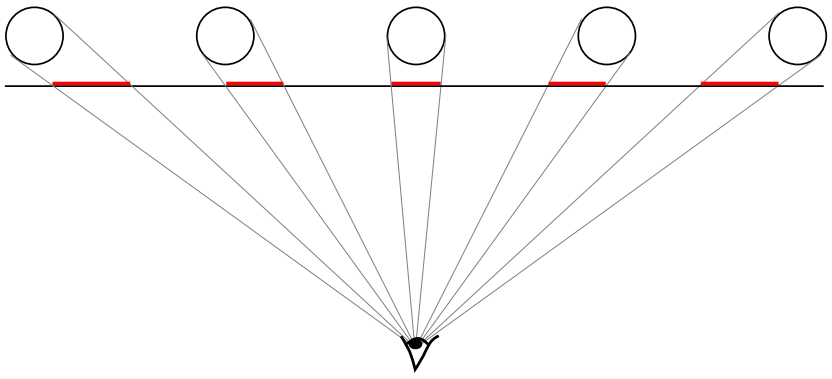


Figure 3.3: A top view of of a row of columns placed parallel to the image plane: their mathematical projection produces perceptual distortion in the picture.

Renaissance paintings avoid these awkward perceptual distortions. Doing so requires perspective practices that take into account more than the geometry of mathematical perspective.

3.1.3 Practical

Renaissance painters were artists, often mathematicians, and were also business men. The clients' expectations were high and the competition fierce. Paintings were commissioned, contracts containing detailed instructions written, often stipulating that the master was to paint the foreground figures, leaving the background to workshop assistants [89, p. 111]. The background contractual emptiness encouraged the master to innovate, especially with the new technique of perspective [89, p. 112–114]. It was essential to find affordable and effective methods for achieving realistic depth. There was much experimentation, followed by refinement of perspective methods. In parallel, measuring tools and mechanical aids [65, Part II] were developed to assist artists creating realistic paintings.

3.1.4 Summary

These three challenges—how to handle depth and image plane composition simultaneously, how to eliminate perceptual distortion, and how to be efficient—were important problems that Renaissance painters faced. Today's image makers face similar challenges. Renaissance paintings provide many examples of addressing them successfully, art historians have examined how it was done. Based on my readings of the art history literature, the next two sections describe the Renaissance approach to realism, which was developed to tackle these challenges.

While reading art history, I was seeking ideas for developing a structural framework that would provide compromises among the following:

- mathematical accuracy,
- perceptual distortion, and
- a realistic simulation of depth in the presence of planar composition.

Despite excellent research on Renaissance paintings and the painters' rare writings, mysteries remain. Thus, my synthesis is partial and approximate. Convenience and practicality, which are subjective, are central to it. My simple synthesis is only a small part of this large topic.

3.2 Planar Composition

The pictorial geometry of Renaissance paintings intermingles planar composition and simulated depth. In this section, I describe aspects of Renaissance practice demonstrating that the pictorial space is organized in the planar surface: proportion relationships and series are used

to size shapes so as to create a geometric pattern on the canvas; lines on the canvas are used to place objects. Composition is based on visual rhythm and balance: geometry on the canvas is the element of control. Understanding the demands of planar composition is essential to understanding Renaissance approach to portraying depth, which is described in the following section.

3.2.1 Proportion Relationships

Renaissance art contains strong visual rhythms, which have classical Greek origins. Using proportions in the design of shapes, from the human body to architectural elements, was common.

Studying the proportions of the human body was prevalent in the Renaissance. The Roman architect, Vitruvius, who studied the body proportion, strongly influenced Renaissance painters. Vitruvius found correspondences between measures of the human body and the geometry of Euclid: 'a square encloses the body while the hands and feet touch a circle with the navel as center. The figure is divided in half at the groin, and by the golden section at the navel' [33, p. 15]. He recommended that architectural design follow such a system of proportions and symmetries, supposing that visual harmony is based on human dimensions. In the Renaissance Leonardo da Vinci and Albrecht Dürer refined Vitruvius's classical proportions of the human body and face [33]. It is thus not surprising that Renaissance paintings contain figures whose proportions guide the planar divisions of the composition as demonstrated in the next section.

Portraying depth brings another kind of proportional relationship, the geometric series that is inherent to perspective. Leonardo and Piero studied its pattern as size diminishes with distance [65, p. 28 and 45]. Figure 3.4 imitates one of Leonardo's study: the diminution doubles as the distance doubles. In effect, discovering that 'the apparent size of the object on the plane is inversely proportional to the distance of the eye to the object' [65, p. 46].

While this observation may seem trivial, it demonstrates the need to reproduce the discrete relations formed by the pattern produced when depth is portrayed. A regular structure like a tiled floor creates a geometric sequence when projected as do all regular intervals in depth: their projections therefore being visually engagement. By studying geometry in the image plane, Renaissance artists detected the visual rhythm of this fundamental law.

Leonardo's simple study of depth diminution shows that an arithmetic sequence projects on the canvas to a geometric sequence. Discovering the rules of perspective, Renaissance artists were interested, beside its accuracy, to portray depth realistically to its geometric pattern on the image plane.

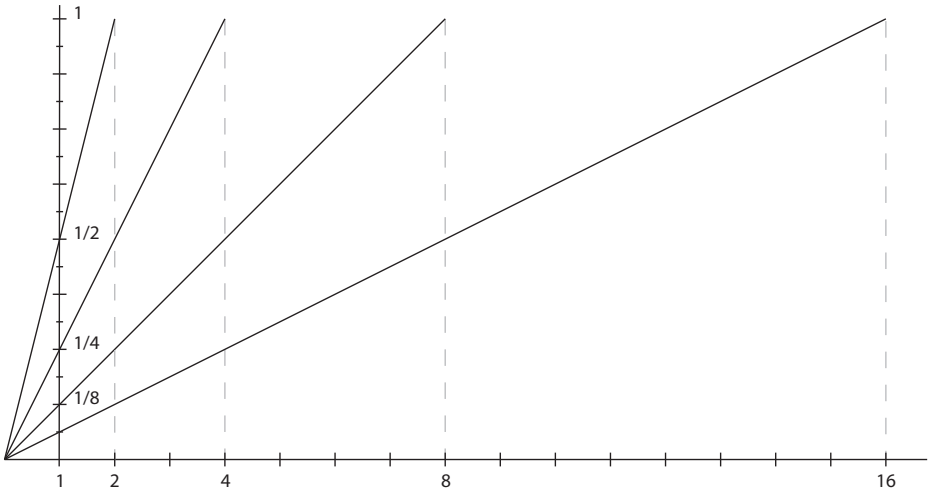


Figure 3.4: Diminution sequence as distance of the object from the eye double. Adaptation from Leonardo diagram study [65, p. 46].

3.2.2 An Early Study Evidence

Renaissance painters used repeated proportions to position important elements of their compositions, thereby creating geometrical relationships in the planar space of the canvas. While it is often hard to know the exact planar patterns that organize a specific painting, I present next some evidence that fixed measures were used to explore and plan the composition to a painting.

An early composition study for *The Entombment* of Raphael, Figure 3.5, includes marks that indicate the use of proportion relations: Thomas uses them to study the geometrical relationship of the composition [104]. On the left side a vertical line is apparent. It is divided into three equal parts and its midpoint is marked. Christ's head is positioned at the height of the midpoint. The length of this line is the horizontal distance separating Christ's head from the right edge of the drawing. Thus, the head resting on Mary's lap was positioned geometrically on the canvas. A horizontal proportion line, labelled with the number two, has its right end-point below the drawing region right corner and its left end-point below the knees of Christ: its length is one third of the drawing width. Finally, a ringed dot is marked on the drawing horizon line, at the eye height level of the standing figures. The lateral position of this dot is close to the center of the drawing width.

As an early study, it contains explicit proof that measured proportions determined positions

Figure 3.5: Early study of *The Entombment* by Raphael (1483–1520): pieta, composition of ten figures (The Ashmolean Museum, Oxford).

of important elements in the composition. Thomas's description is objectively supported by guides visible in this early study. It is worth noticing that the proportions used are one-half and one-third. It is convenient to rely on small integer fractions: dividing by two, three or four is easy and quite precisely evaluated by eye, without measuring. (Beyond four, free-hand evaluation is tricky.) Small integer fractions recur in Renaissance canvas geometry and object proportion.

Next, the lines on the canvas that organize the compositions of two well-known Renaissance paintings are described: *The Baptism of Christ* and *The Ambassadors*. In effect, contemporary art experts explain their compositions by geometric relations in the canvas plane.

3.2.3 Painting Examples

2D Geometric Pattern. From the early Italian Renaissance, *The Baptism of Christ* (c. 1450) by Piero della Francesca is shown in Figure 3.6 (a), overlaid by the geometrical figures that describe its composition. A 2D shape, an equilateral triangle—with its reflection lines and its inscribed circle—is the skeleton that controls the positions of important image features. The perspective in the landscape forms should not be lightly considered: its coherence for the unified outdoor scene illustrates Piero's expertise of linear perspective [69].

(a) (b)

Figure 3.6: (a) *The Baptism of Christ* (c. 1450) by Piero della Francesca © The National Gallery, London. Overlaid is the lines suggested by Carter that organizes the painting as a 2D composition [69]. (b) *The Ambassadors* (1533) by Hans Holbein the Younger, © The National Gallery, London. In the foreground is an anamorphic projection of a skull. Overlaid lines indicate, looking down, where to stand to see the skull, and looking up toward the face of Christ on the crucifix, the eye of an ambassador and significant points on the astronomical instruments [81].

Art historians are convinced that the overall compositional organization is two-dimensional, and not three-dimensional, and that significant elements of the painting are aligned with geometric patterns in the painting [15, 34, 53, 69, 81]. The painting is a semicircle surmounting a rectangle. This shape has a prominent vertical mid-line, which coincided with the mid-line of the church. The mid-line a stable axis in the painting, on which Christ is centered. The lower portion of a circle completing the upper semicircle, the dashed circle in Figure 3.6 (a), runs along several prominent lines in the painting: the top of Christ's loincloth, John the Baptist's left arm, and the shirt the bather is pulling over his head[1]. It separates the spiritual upper parts of Christ, John the Baptist and the angels from their worldly lower parts.

In addition to the circle a base up equilateral triangle is the main compositional structure of the painting. Features of the triangle connect many significant elements of the image, a few of which are sketched next.

[1]It crosses the angels over diagonals of their upper bodies: stomach, chest and head. The circle symmetrically closes curves on the angels, such as the cloth fold of the angel in the center and the hair line of the one on the left. The directions of their hands reinforce the curvature.

- The base is the horizontal diameter of the circle, and the ends of the base are at the edges of the painting.
- The base separates the semi-circular upper element from the rectangular lower part.
- The open wings of the dove, the Holy Spirit, lie along the base; the dove itself is at the centre of the base.
- The hands of Christ are at the centre of the triangle.
- The toe of Christ's right foot is at the lower apex.
- The elbow of John the Baptist on Christ's left is at the centre of the triangle right edge.
- The elbow of the middle angel on Christ's right is at the centre of the triangle left edge[2].
- The vertical mid-line of the triangle, along which Christ stands, divides the rectangular part of the image into two halves. To Christ's right are spiritual elements, the group of angels and the Tree of Life; to Christ's left are earthly elements, John the Baptist and the naked bather.
- Only the Tree of Life extends into the semi-circular part of the image, the realm of the Holy Spirit.

The equilateral triangle organizes the composition, pulling together the main elements. As a symbolize for the Trinity, it provides a cultural association that reinforces the picture narrative. The triangle divides the canvas in many ways, by its vertical axis, its horizontal edge, the midpoints of its edges and its inside versus outside. The ubiquity of geometry controlled by the equilateral triangle demonstrates that Piero used 2D geometry to locate the picture elements.

The Baptism's composition was conceived in two dimensions, which is confirmed by Piero's planning the painting. First, he blocked in Christ and John the Baptist. Then he painted the landscape, after which the foreground figures were added: for example, most of the foreground angels were painted over the landscape. Following typical Renaissance practice, the foreground figures were probably drawn separately, from life, as cartoons, which were later transferred to the painting.

The Baptism of Christ is not a random snapshot of a naturalistic scene: each element contributes to the overall harmony of the picture, largely by creating conspicuous planar relations.

2D Alignment. From the late Northern Renaissance, *The Ambassadors* (1533) by Hans Holbein the Younger is shown in Figure 3.6 (b). An unusual object superimposed on the scene, an

[2]Examination of the underpainting shows that the group of angels were moved to create this relationship.

anamorphic projection of a skull, seems unrelated to the painting. Behind it, the scene realistically portrays the eponymous ambassadors and two shelves full of objects, mostly scientific instruments[3]. The skull is considered by most art historians to be essential for understanding the complex symbolism Holbein included in the painting.

The anamorphic skull in the foreground of the painting enforces an unexpected geometric reading of the picture as North explains [81]. When the painting is observed from a secondary, unconventional viewpoint, two problems dissolve: the mysterious object reveals itself to be a skull, and reveals an alignment of objects that tells the story of the painting. The anamorphic skull is recognizable as a skull only when seen from a very particular station point, near to the picture plane, from just right of the frame and at the sixteenth century eye level: the intersection of the two lines overlaid on Figure 3.6 (b) is this secondary viewpoint.

From this secondary viewpoint, the viewer can look down to the skull, or up at the same angle ($\arctan(1/2)$) to a crucifix, following the overlaid lines. The view to the crucifix passes precisely through significant objects in the painting: the eyes of Christ, the left eye of the ambassador, the center of the globe and so on, symbolically telling the story of Good Friday. This interpretation is reinforced by shadows on the astronomical instruments which give the precise time, 4pm on Good Friday 1533, which is according to Medieval tradition, exactly a millennium and a half after the crucifixion.

In addition to this basic plan North finds other significant patterns in the frontal plane, of which a few are listed below.

- A plumb line in the exact mid-line of the picture.
- An astrological chart with significant items in the painting indicating the positions of the planets at the time of the painting.
- Two identical hexagons, one in perspective on the floor, one in the frontal plane.

Holbein has produced a painting with patterns in the picture plane that are essential to understanding it. Yet perspective is not neglected. For example, the anamorphic skull demonstrates mastery of perspective. Leonardo investigated anamorphic projection [65, p. 50], but the first treatise presenting its practice was only published in 1693 by Andrea Pozzo long after *The Ambassadors*, so that Holbein's skull is an important discovery [74]. In addition, the pattern on the floor and all the individual items in the picture are painted in precise perspective.

How does Holbein satisfy all the constraints required by these many patterns? In the next section, I return to the application of perspective that gives him the freedom to create planar composition, combined with a realistic portrait.

[3]The upper shelf holds astronomical and geometric instruments; the lower arithmetic texts and musical objects.

3.3 Artificial Perspective

In late medieval and early modern times—a period known since the mid-nineteenth century as the Renaissance—a new method for portraying three dimensions was (re)-discovered, linear perspective. Since its elaboration in Renaissance painting to solve practical problems of artistic optics, it has been developed into the science of geometrical optics, and formalized as projective geometry. Today, having both geometric and algebraic models of linear perspective, it is common to treat Renaissance perspective as an imperfect version of modern theory. Yet art historians assure us that this description understates the richness and flexibility of artistic practice in the Renaissance. Renaissance artists used perspective not only for producing the effect of a unified space in which objects exist, but also for enriching the depiction of individual objects. The artificial perspective of the Renaissance painters does not adhere rigidly to mathematics. 'Violations of perspective' in Renaissance art are not apologetic expedients, introduced to avoid other distortions, but a demonstration of the primality of the objects depicted, objects that provide the image elements of which the image is composed.

Renaissance practices deal with the 'pictorial' nature of realistic representation, illusory depth agitating the picture plane while respecting its physical two dimensional nature. Indeed, Renaissance perspective is a judicious adaptation of depth depiction in the image plane. In this section, I argue that two aspects of Renaissance perspective distinguish it from the perspective model of computer graphics.

1. Renaissance painters develop 3D using construction lines directly on the image plane. Using them to depict a tiled floor in perspective collapses the 3D problem of managing depth in the scene to a 2D one on the canvas. Its structure permits to construct a unified consistent space.

2. Simultaneously, Renaissance painters experience the limitations of unitary perspective and go beyond it. Their application of perspective is coherent but 'pluralist': local perspectives are precise, perceptual and practical.

For each aspect, supportive constituents are presented before describing their illustration in two paintings. In effect, Renaissance practice with respect to perspective is a combination of global realism often formed by architectural structures, and of local realism in the foreground volumes, mostly represented by human figures involved in the narrative of the painting.

3.3.1 A Global Space

Renaissance artists were the first painters to present a clear sense of depth using linear perspective. This subsection briefly explains the basic constituents used to do so and describes two exemplary results of their application.

Construction Line Methods. Renaissance artists developed and used many methods for perspective [65, 36]. Some were designed for depicting geometric shapes. Architectural forms are the most common example. Their volumes, delimited by rectilinear lines, are 'amenable to geometric treatment' [65, p. 32].

Construction lines are lines artists draw on the canvas to guide the process of flattening the 3D world onto the image plane. Most are invisible in the final picture, but they are essential for positioning, aligning, and sizing objects portrayed in the picture. They are a tool for geometrically evaluating on the canvas the locations of 3D objects. With them artists draw regular shapes in perspective or evaluate the space enclosed by a set of planes directly on the canvas. Evaluations are made by the intersections of construction lines.

Sets of construction line evaluations form methods. The methods, which are essential to Renaissance art and thus fundamental to my framework, are presented in Chapter 4. Chapter 5 then turns those illustrative methods into geometric algorithms for my framework. A synthesis of Renaissance construction line methods is the traditional way that artists develop the geometry of perspective on the canvas, using correspondences between the line directions that are relevant to rigid objects in the scene. Remarkably, Renaissance construction lines project 3D geometric relationships to geometrical constructions in 2D. The new approach to perspective in computer graphics proposed in this thesis is an expression of this capability.

The Tiled Floor. Geometric realism depends on coherence. A unified space is coherent. A unified space depends on a global structure: the tiled floor in perspective is one such structure. The tiled floor in perspective depends on selecting the position of the viewer and its view direction; alternatively the appearance of the tiled floor indicates the viewer's position and his line of sight.

Leon Battista Alberti in his treatise *On Painting* (1436) [3] describes a construction line method for creating a tiled floor in perspective [65, p. 21–23]. The construction of the tiled floor is fundamental: Chapter 4 examines it in detail. In effect, the tiled floor is the global binding structure to which the shapes of the picture adhere. Its structure, squares repeated periodically,

defines the configuration upon which the scene portrayed is rooted, as it discretizes the ground plane in a regular grid. It is the critical reference for constructing other geometrical shapes. Construction lines that evaluate locations in the scene depend on the tiled floor appearance: the objects are sized in relation to it and arranged above it.

(a) (b)

Figure 3.7: (a) *Saint Jerome in His Study* (c. 1510) by Vincenzo Catena, © The National Gallery, London. This painting is an example of a rigid application of linear perspective. (b) *The School of Athens* (c. 1510) by Raphael is a collage of constructions: construction lines indicating that the architecture adheres to a principal perspective; insert highlighting the individual perspectives for the off-center spheres.

Examples. Construction line methods, especially the tiled floor construction, provide the unified space of Renaissance paintings. The geometry of linear perspective is present in the representation of the 3D volume in which the scene space is situated. In other words, the architecture presents the principal perspective of the picture to the viewer.

For example, Vincenzo Catena's *Saint Jerome in His Study*, shown in Figure 3.7 (a), is a fully unified space. Its perspective is defined by a simple box, which unifies the picture by making explicit the architecture of the room. The beholder's feeling is at 'its calmest and most lucid' [36, p. 125]. The open doors on the back wall are necessary to slightly disturb the overly restful scene. If they were too close, they would overload the frontal exposition of the scene, strongly reinforced by the frontal planes of Saint Jerome's desk and the back wall. Instead, they provide an unexpected balance by reinforcing the too quiet longitudinal directions, among which the discreet Christ's cross horizontal bar.

Similarly in *The School of Athens* of Raphael, Figure 3.7 (b), the overall architecture forms an accurately mathematical 3D space unified by a rigidly linear principal perspective. The perspective structure of a tiled floor is apparent. In fact, Raphael employed many assistants who carefully measured the architecture that defines the space occupied by the philosophers. Hundreds of three-dimensional measurements were made, perspective calculations were performed, and the results transferred to the surface on which the picture was created [36, p 227–228]. The architecture is geometrically accurate to a high-degree of fidelity. But as we will see in the next section, there is more to it: why are the lateral spheres represented as circles? This question has intrigued many computer graphics researchers [113].

3.3.2 Local Collage of Constructions

It is simplistic to consider Renaissance perspective as a mathematical application of projection from the 3D scene to the 2D image plane. The art historian James Elkins, in his book *The Poetics of Perspective*, challenges the anachronistic view that perspective as a rigid process, to which Renaissance paintings owe their unified appearance [36]. Drawing on Renaissance texts and paintings, he demonstrates that writers and artists had a 'pluralist' approach to perspective, that many 'compatible perspectives' co-exist within a single realistic picture, in contrast to the 'monolithic, mathematical' perspective of modern rendering [36, p. xi]. In particular, Elkins draws heavily from the main source on Renaissance perspective, the book *Vite de più eccelentri pittori, scultori, ed. architettori* (1550) by Vasari [106], an authority on Italian Renaissance art. Elkins describes a rapid evolution from pictures that unify space by a unique perspective to pictures that show objects possessing individual, local perspectives.

Indeed, Renaissance paintings having a single central perspective are rare. Early in the Renaissance the calm atmosphere of central perspective is disturbed. As a composition is subject to many constraints: '[t]he problem of projecting a 3D scene onto a 2D picture is in fact overconstrained[: it] is not possible to satisfy all the "desirable" properties' [29]. In Elkins's words, the artist must 'sort out the objects in a perspective picture so that are all fully visible from a single place, one must rearrange them in two ways. First, movable objects need to be separated; [...] Second, the perspective box must be arranged so that receding lines do not occlude one another at odd angles' [36, p.125–126]. At first artists accepted these limitations to obtain clarity, separating objects and arranging architecture: the size and location of objects constrained, the mood peaceful, as shown in Catena's painting (Figure 3.7 (a)). Later, Renaissance artists experimented with other strategies, such as combinations of perspective constructions, thus stepping outside the metaphor of a single window through which the world is seen.

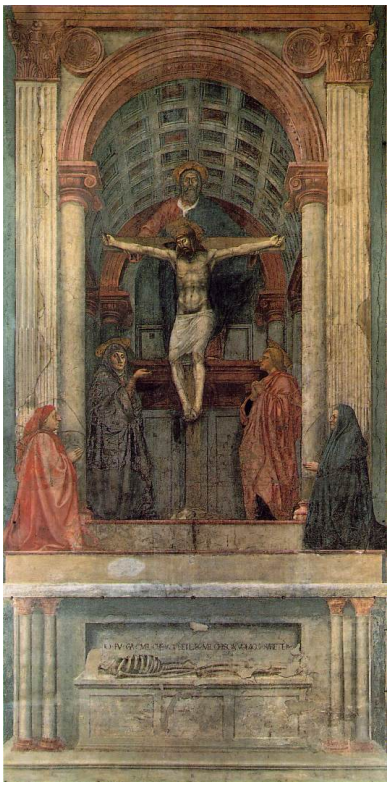


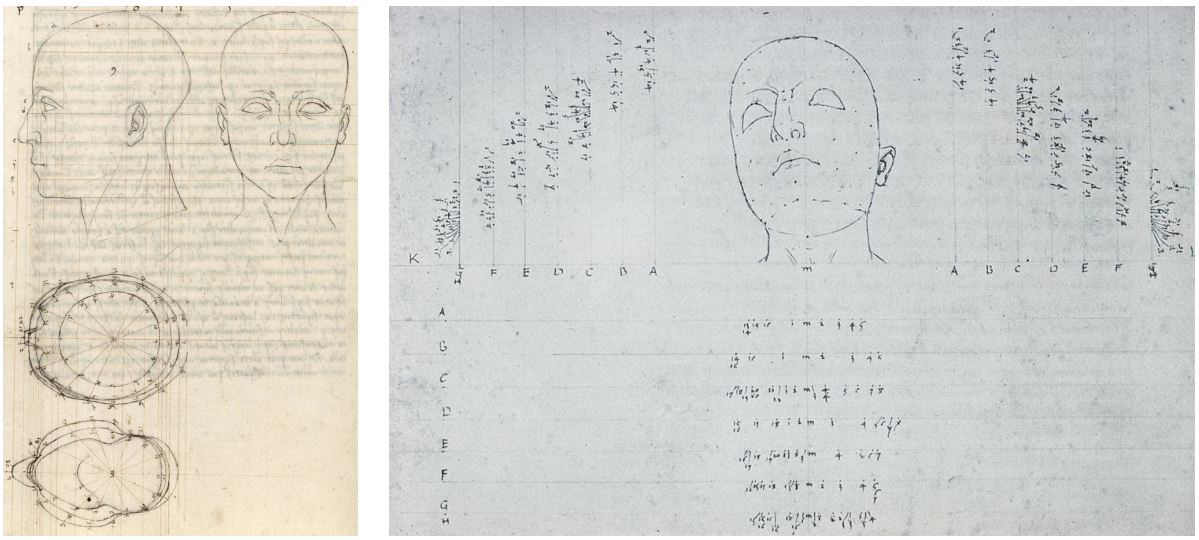
Figure 3.8: *Trinity* (c. 1426) by Masaccio. One of the earliest paintings that gives the illusion to be unified by perspective. The 'sheer sophistication of the construction' [65, p. 21] makes it 'extremely naturalistic' compared to its predecessors [37, p.245]. Vasari instead was first arrested by the figures.

Praise for Individual Perspectives. Vasari's comments on pictures emphasize individual perspectives: he ignores the overall perspective of paintings, while praising earlier for local perspectives applied to individual objects, a chair, a horse painted in perspective by Piero, and a barrel vault by Masaccio [36, p. 53–56]. 'For Vasari, good paintings could be "full of perspectives"', [36, p. 55]. Thus Elkins concludes that Renaissance art focuses on objects and space is secondary. Realism has an 'object oriented' focus, in contrast with the modern approach, which is 'space oriented' [36, p. 15].

For example, discussing the important work of Masaccio *Trinity* (c. 1426) (Figure 3.8), which opened a 'new kind of perspectival rigour' [65, p. 21] Vasari praises first the beauty of the figures, and only then the accomplishment of 'a barrel vault drawn in perspective' [36, p. 53]. The two accomplishments are clearly differentiated. The second one concerns perspective but as a part of a picture rather than as the unifying space bounded by the vault [36, p. 53].

This account distinguishes the figures, which are only loosely engaged with the architecture construction, revealing the essence of Renaissance perspective. The goal is the figures in relief, 'the ultimate aim of painting' [89, p. 119]. The combination of colour, light and shade were most important for giving *rilievo* (volume) to the figure, the roundness enhancing the illusion of lifelikeness [89, p. 180]. Thus, we now discuss the geometry of figures.

Human Figures. Human figures are the usual focus of paintings in the Italian Renaissance. They are actors in the story of the painting. They are portrayed either in front of a 3D scene, in the foreground or within a 3D scene, in the mid-ground. Human bodies and their clothing are geometrically intractable compared to architecture [65, p. 32]. Nonetheless, Renaissance painters precisely represented them in perspective. Piero in the third book of *De Prospectiva pingendi* [25] discusses human shapes, devising elaborate methods for studying them perspectivally (Figure 3.9). His methods rely on geometrical calculation to find intersection points between plan and elevation views [65, p. 32–34][37, p. 102–112]. Piero's studies of human heads are particularly interesting because they can be compared to those in his actual frescos.



(a) (b)

Figure 3.9: Two studies of Piero to draw human heads. (a) shows how to draw one that frontally faces the line of sight: on the left column are a side view and two top ones (one from above and one from below); the numbers on their vertical, horizontal and radial axis determine the head construction on the right. (b) shows how to draw a head slightly turned and tilted backward seen from a viewpoint on the vertical axis but below: the lower surface of the chin is visible.

The head study of Figure 3.9 (b) appears in Piero's *Resurrection of Christ*, (c. 1465) (Figure 3.10). The sleeping soldier on the left of Christ has his head at such a tilt and is seen from below. Each soldier has his head tilted in at least one plane, and is seen from below as is the tomb. Christ's head, in contrast, is directly frontal and unforeshortened with his foot seen from above. The differing eye levels are intentional: they separate the riveting head of Christ from the soldiers' heads. While several viewpoints are used, the 3D scene represented is visually coherent. The painting's naturalism resides in its parts, and not in a unifying whole, which is appropriate to the religious narrative. The fresco is large and seen from below [37, p. 109]: the soldier's horizon is appropriately that of the human viewer, while Christ's higher horizon is for a superior realm.

Figure 3.10: *The Resurrection of Christ* (fresco c. 1465) by Piero della Francesca.

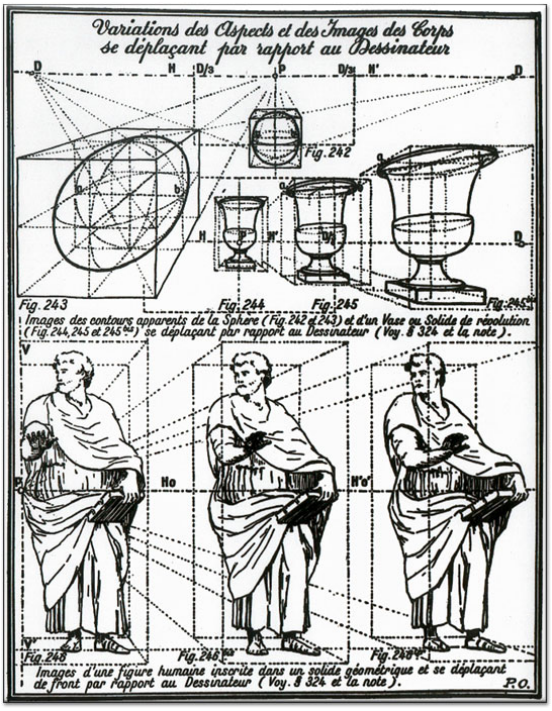


Figure 3.11: Figures from *Perspective Artistique* (1948) by Pierre Olmer [82].

Local Perspectives. For depicting organic shapes locality is essential: the perspective is executed with respect to central axes of the individual object. The local horizon line of the object determines its upper and lower visible surfaces, the horizontal offset position of the viewer with respect to the vertical central axis of the object determines the visibility of its lateral surfaces. The form is drawn so that its three-dimensionality is precisely captured within its local reference frame. In Renaissance paintings, objects rendered with respect to their central viewpoints are recurrent: transferred from studies done in studio or sometimes possibly achieved 'freehand' [65, p. 34].

This choice of local viewpoints for those organic shapes is intentional and not merely an expedient. Renaissance painters used the perceptual benefit of local realism. Indeed, Kubovy [68, p. 116, 125–126] noted that human figures off the principal axis of a central projection appear distorted when drawn in perspective, as do spheres and cylinders: Olmer's diagrams (Figure 3.11) illustrate it. In effect, Kubovy recommends following the Renaissance practice of drawing each organic body—humans, animals and plants—from a center of projection directly centered on it.

The observer looking at a picture from its station point[4] with the view direction perpendicular to the image plane, sees objects laterally far from the line of sight in the periphery of vision and therefore defocussed, so the distortions are invisible. But when a viewer wants to see them clearly he changes position to the object's local viewpoint imaging the object on the fovea. The projection is then bilaterally symmetric, without oblation. Using the object's local perspective, the depiction is perceptually robust, against changes in viewpoint and may be placed anywhere in the final composition. In contrast, under an unitary application of perspective a rounded object placed laterally away from the line of sight appears deformed when closely inspected. That is the advantage of artificial perspective, which is beyond mathematical perspective derived from optics laws, where the eye is modelled as a static organ.

Renaissance painters chose the 'drawing' viewpoint to reveal the object to viewers at all viewpoints. Consequently, many local viewpoints integrate seamlessly into a realistic painting. As Renaissance pictures are large, especially altarpieces and wall-paintings, it was critical for their painters to support distant fixed viewpoint and close observation of an individual object at the same time. Renaissance intentions were not those of abstract art, as their images are not geometric puzzles challenging the camera view: the goal was to represent perceptually on a flat surface the world in front of the viewer's eyes.

Collage. Renaissance artists had many methods for drawing in perspective. Several methods were combined in one painting. Elkins calls this phenomenon 'a collage of different constructions' [36, p. 131]. Renaissance perspective practice is a mixture of 'convention, improvisation, local exigency, and aesthetic decisions' complementing 'mathematical necessity' [36, p. 115]. The artists knew there was no single perfect method for transcribing reality.

The main volume is often drawn in a rigid unitary perspective while local objects are portrayed with independent perspectives, allowing an artist the freedom to achieve detailed representations. The mixture of constructions enables three things:

1. favouring the planar composition over consistency in projection,
2. enhancing realism by including complex local detail, and
3. taking active perception into account, specifically the viewer's varying station point.

[4]Also viewpoint, is 'the artist's location when viewing the object. A drawing, to be consistent, is made as if the station point were a fixed location and the artist's eyes were kept in the same position—without moving from side to side or up or down—throughout the making of the drawing' [7].

Two examples of collages of constructions in Renaissance paintings are now examined. They have detailed objects, each having its own perspective. How are they created? Each merges differently treated shapes—geometric and organic—into a globally coherent 3D space.

Examples. *The School of Athens* of Raphael, previously shown (Figure 3.7 (b)), is known to be an explicit collage of constructions. While the overall architecture forms an accurate 3D space, foreground figures are portrayed locally, each having its own individual viewpoints, none of which is the perspective of the architecture. The result is not a careless attempt at mathematical projection, but a considered combination of perception and creating a unitary architectural perspective.

Indeed, the full scale cartoon (Figure 3.12 (a)), which is the pattern pricked onto the fresco, shows the foreground and mid-ground figures to be a collage of separately executed drawings. The full scale cartoon combined many sheets of paper, each containing an individual cartoon of a single figure or a small group. Each was drawn independently, with its own centre of projection. Each group, which is a few figures located together and at the same distance from the viewer, is a local frontal plane. The round bodies of the foreground figures are curved volumes, projected onto the image plane, as though the fresco's station point faces the group. As such they are perceptually better representations than if they had followed the perspective of the scene architecture, which would have induced lateral deformations (Figure 3.11[5]).

(a) (b)

Figure 3.12: (a) Cartoon of Raphael's *School of Athens* used to prick the fresco: assembled from many separate sheets of paper, each representing a group. It sets the composition of the foreground figures. (b) Detail from *The Ambassadors*, © The National Gallery, London. This detail 'shows what seems to be a collage construction' [55, p. 100]. The vanishing points of the two books are at different heights: they were portrayed from different viewpoints, which goes unnoticed.

[5]Olmer's study of lateral distortions actually uses Aristotle, one of the two central philosophers from the *School of Athens*.

Raphael could decompose the foreground and mid-ground figures into many independent studies because they group into distinct frontal planes. The figures in a frontal plane are the same distance behind the image plane. The planes overlap but do not intersect in depth. Thus, they are easily assembled, at sizes appropriate to their depth in the painting. Separation of frontal planes was routine in Renaissance perspective practice [28, p. 41]. Indeed, it is possible to decompose the figures of *The Baptism of Christ* into four frontal planes: two in the foreground, one containing Christ and Saint John, the other with the three angels; and two in the mid-ground, one with the bather, the farthest one with the group of Magi. In effect, Arnheim noticed that in central perspective 'objects are placed frontally whenever possible' and should not be deformed unless the 'task of representing depth requires it' [5, p. 286]. Raphael's treatment of the foreground figures, for which the full scale cartoon has survived, follows Arnheim's rule. For architecture, depth representation requires that the longitudinal planes that enclose the space are directed to the vanishing point of the picture central perspective: on the horizon facing the picture station point. Thus, the architecture is not part of the collage of frontal planes [89, p. 116]. The full frontal cartoon fits into the architectural space of the painting.

Raphael achieved successfully this further collage of integrating the foreground collage with the central perspective of the architecture: the flexible scholars inhabit unyielding stone. The organization of the philosophers depends on easy perception of their scene depth. The collage anchors each philosopher either to a stair or to the ground. In the image plane they largely occupy the foreground area. The arrangement in depth is unmistakable. Each group is firmly fixed in architectural depth, rooted on steps or anchored to the tiled floor, standing, kneeling or lying. Thus, the figures form groups and the groups form 2D patterns that are complex and arresting. Their sizes of figures are consistent with the scene perspective. The architecture ground geometry—the tiled floor and the steps—giving a structure to establish the scale for foreshortening the figures on the full cartoon according to their positions in the picture depth.

Similarly, *The Ambassadors* of Hans Holbein the Younger (Figure 3.6 (b)) contains many local perspectives. 2D patterns govern the positions of the objects on the canvas (§ 3.2.3). To satisfy all the constraints implied by these patterns Holbein's perspective is, in the words of Elkins, 'only gently mathematical', but imperceptibly so.

The sizes of objects are adjusted to conform to 2D patterns. That is, each object in the picture is painted precisely in its own perspective, and the objects are arranged in a collage of perspectives. While each object has a convincing individual perspective, irrelevant inconsistencies among the different local perspectives can be scrutinized. Hockney, for example, pointed

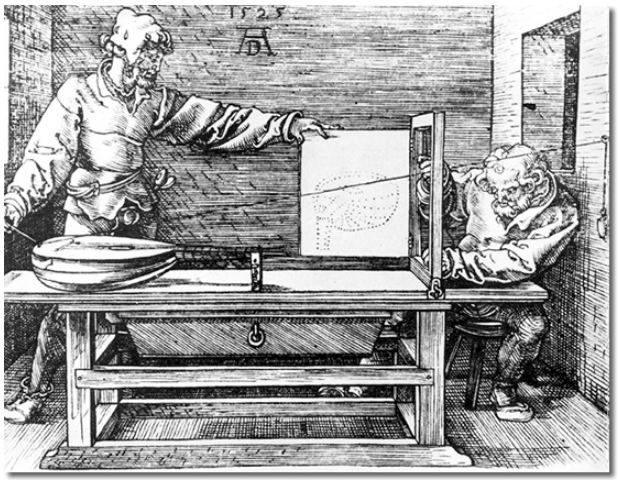


Figure 3.13: An engraving by Albretch Dürer (c. 1525) illustrating the use of a *portillon* to study the perspective of a single complex shape: a lute.

out the non-alignment of the viewpoint used to portrayed the two books posed on the same shelf of the stand [55, p. 100] (Figure 3.12 (b)). The strong parallel lines provide unmistakable individual perspectives, which are indeed 'object oriented', rather than 'space oriented'; planar composition overrules depth consistency. The deviation from the rigid application of unitary perspective goes unnoticed. Most important are the books' poses and expressions: one book is wide open with the music score readable; the other, slightly-opened, is an identifiable arithmetic text. Each is oriented to be recognizable and convincing as a particular book—closed books do not reveal their content. Similarly, the lute is drawn to expose its essential feature, down to the broken string. The three-quarter view from above is optimally recognizable; Albretch Dürer used the same view in his woodcut (Figure 3.13) illustrating a technical aid used to study the foreshortening and flattening of a single object.

In *The Ambassadors*, each object has a perspective based on a viewpoint more or less in front of the object. Each possesses its own horizon line and principal vanishing point, not yielding to unified perspective. Holbein expected the viewer to be active while observing the painting, moving to observe each object from the best vantage point, and the object is painted for such viewing. (The extreme view point required by the anamorphic skull is the most obvious example of this expectation.) The collage technique is well adapted to the demands of painting in oil. Paintings took months to complete, with the models, animate and inanimate, available for only short times. Objects were either drawn from life and transferred to the painting using a cartoon, or painted individually with the other the objects not present. Fortunately, North was able to identify the originals of many objects in *The Ambassadors*, making possible quantitative

reconstruction of each perspective geometry [81]. He found that most objects had been scaled to conform with patterns in the picture plane. Scaling is harmless, since most viewers do not know the size of the original objects. The painting has simple occlusion relationships, which simplify creating the painting from parts. This property it shares with Piero's *Baptism*.

These two examples show the richness and complexity of Renaissance perspectives. There are many other examples [36, 89]. Beauty and balance in the picture plane rarely can be created without rearranging the visible world. Renaissance paintings are convincingly realistic, yet contain patterns in the picture plane and in depth.

3.4 Conclusion

Having discussed the creation of spatiality and geometry in Renaissance pictures, looking for new ways of creating realistic images with the computer, I conclude with the main points that are important for my framework.

- Frontal plane relationships among foreground elements contributes to the story of the painting.
- Perspective is applied element by element, to show each in its most characteristic and legible view.
- Nature/architecture creates a perspective space in which foreground elements are situated.
- A tiled floor structure helps to simulate the picture unifying depth and permits to locate and size elements throughout it.
- Construction lines are used to get architectural perspective true.

Indeed, artists work from lines on the canvas that mark significant locations for their composition and for the simulated depth of the picture. They arrange natural elements, including humans, in a collage, without harming identity or legibility; occlusion relationships among the elements are commonly simple, making separation of frontal plane possible. Within generous limits element by element variation in perspective is unnoticed by the human visual system and on in many occasions, independent perspectives are preferred locally.

Architecture needs a unifying perspective, a single projection that provide a space in which the narrative foreground takes place. Built from plans or measurements, it is drawn using

2D construction lines, which are 3D alignments and intersections projected, in the sense of projective geometry, onto the canvas. Methods, to be discussed in future chapters, allow the artist to build geometric shapes directly on the canvas, instead of building 3D versions of them, which are then projected onto the image plane. The construction of the tiled floor indicates the correct foreshortening and concrete placement for elements by the tile size and location at discrete depths: it functions as a 'compositional device' [31]. Artists commonly use central perspective[6], a simple perspective configuration for the scene architecture, which does not obstruct the planar harmony of the picture or distract attention from the foreground elements.

This chapter, and those main points, do not form an algorithm for creating a picture; incomplete heuristics were brought to light. What is not specified is supplied by the artist's tacit knowledge. To implement those main points on a computer, which might relieve the artist of the drudging of drawing many construction lines, they must be formalized. Using an essential illustrative set of constructions and a representation to include cartoons effectively is the solution proposed in this thesis.

For example, artists' cartoons are replaced in my framework by panels: picture within a picture, which have 3D properties such as depth and orientation. An element projected once onto the panel as picture plane, is projected a second time onto the picture plane of the picture of which it is an element. It is not the recursion that is fundamental to the panel abstraction, but the broken perspective coherence: standing at the station point of a picture and looking straight, the picture perspective and the perspective of its environment coincide, but moving in position or gaze direction breaks it.

[6]Today we call it one-point perspective.

Chapter 4

Construction Line Methods

This chapter presents construction line methods that are used in my framework, first as a set of computer-simulated straight-edge and compass constructions (Chapter 5), then implemented as 2D projective transformations that build into the artist's constructions (Chapter 6).

Most construction line methods were developed in the Renaissance: those for projected shadows were discovered later, possibly as late as the nineteen century. I introduce them independently of my framework to emphasize their origins and to illustrate the interesting double projection problem they solve.

In particular, this chapter focusses on the construction of the tiled floor. Developing methods for projecting tiled floors taught Renaissance artists the properties of geometric mapping: between the 3D space of the scene, and the 2D space of the canvas surface. In effect, studying the appearance of the tiled floor in perspective, Renaissance artists discovered how the 3D problem of projecting a scene onto the image plane can be transformed into finding 2D intersections on the canvas. Examples of other constructions to depict volumes and shadows are reviewed briefly because those construction line methods are described as algorithms in Chapter 5.

4.1 Introduction

Much is known about the perspective methods of the Renaissance, through art history [37, 65], histories of Renaissance mathematics [4], and Renaissance accounts [3, 106, 25]. These texts demonstrate that early in the Italian Renaissance methods were known to draw in perspective tiled floors circles, realistic arches and so on. Many variations exist of these methods [65].

The contributions of this thesis are based on this scholarship. This chapter uses it to identify the concepts needed to simulate 3D geometry on the image plane. These concepts, many of which present in the construction of the tiled floor (Section 4.2), suggest a different way of creating the appearance of 3D in computer graphics. This chapter emphasizes concepts revealed in artistic practice, hitherto bypassed in computer depiction.

4.1.1 Connection of Spaces

Current ideas of perspective take projective geometry for granted. Understanding it is unnecessary for writing code that projects correctly from 3D to 2D. Few programmers create it by hand, as could most Renaissance artists. Perspective is now an 'out of sight' technology, over which users have only loose control: the viewer tends to be a 'servant' of perspective rather than its 'master' [32, p. 58].

Renaissance artists experimenting with perspective using a straightedge and compass did not have any explicit recipe. They progressively developed intuition and practice, beginning with the perspective of a simple scene viewed in central perspective (Chapter 3). They sought 2D patterns that gave the impression of depth, from which they generalized rules for simulating perspective on a surface.

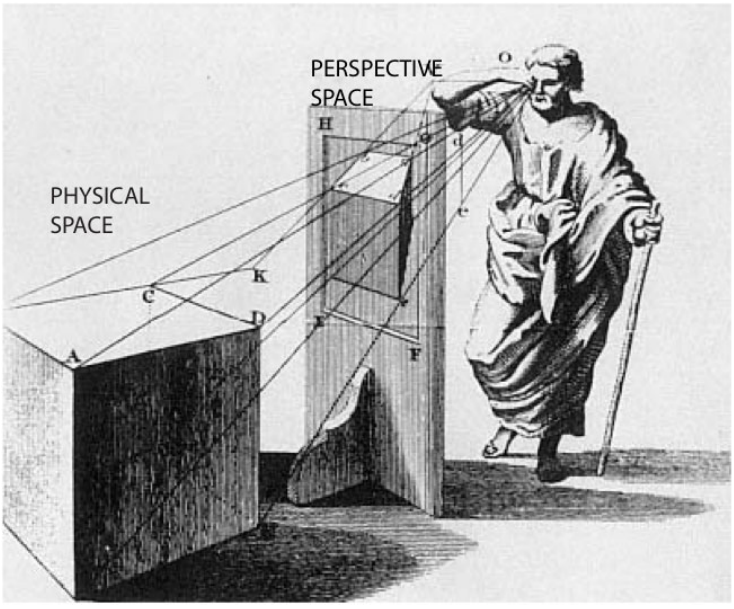


Figure 4.1: Illustration of the relation between the physical space all around the image plane and of the perspective space, which is only the canvas surface.

A painting is done on a 2D canvas; its content is a view through the 'transparent' canvas of a 3D scene. Thus, two spaces interact. The 3D scene that is the *physical space* with 3D Euclidean geometry. The viewer and artist are also located in the physical space. The canvas is a *perspective space*, which has 2D projective geometry. Figure 4.1 illustrates the two spaces. While visualizing the relationship between the two spaces is easy, formalizing an interaction rich enough for artists is not simple. The window metaphor describes the geometric configuration, but does not provide a practical straightedge and compass solution.

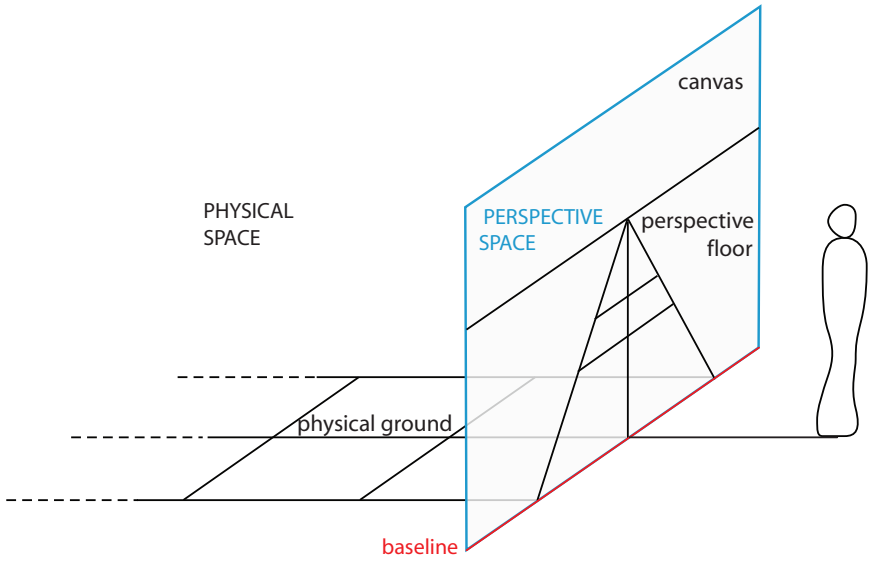


Figure 4.2: The physical ground behind the canvas: there is no gap between the canvas lower edge and the physical space behind it. the baseline touches the physical ground.

In a reversal of computer graphics, painters choose first the view. In addition early Renaissance artists chose a specific simple configuration connecting the two spaces. Their paintings, commonly frescos or altarpieces[1], were large, with life size foreground objects and the canvas extending to the floor, the physical ground plane. The horizon would be at eye level and the ground behind the picture was fully visible on the painting (Figure 4.2), in contrast with easel painting (the canvas in Figure 4.1). This canvas geometry reveals a fundamental connection between the two spaces: the perspective ground at the painting's lower edge intersects the physical ground. Thus, the spaces coincide at the canvas baseline where the artist can transfer

[1]The statement that follows is not true for easel painting. Easel paintings were used for head portraits, but were not common until late in the Renaissance. Those portraits do not contain much perspective scene. Most common were human figures in full portrait in Renaissance perspective scenes.

lengths from one space to the other. In effect, Renaissance frescos see the 3D world through a French door; easel paintings through a window. This description is simple but conceptual; it describes most Renaissance paintings, modulo an uniform scaling [4, p. 334–336].

In this chapter, I explain this strong connection between the two spaces, which helped Renaissance artists to solve the challenge of depicting 3D. Evaluations of construction line intersections transfer sizes between the two spaces, effectively simulating 3D objects using only 2D geometric constructions.

4.2 The Tiled Floor

On the canvas, a perspective tiled floor is the reference for scene depth. It establishes the foreshortening of depth recession. The tiled floor determines the view geometry and painters began by drawing a tiled floor, even when would be invisible in the final rendering.

In the physical space, a tiled floor is an array of squares on the ground plane, a geometric grid perpendicular to the image plane. In this configuration, central perspective, the floor is a simple, intuitive depth grid. The parallel horizontals contrast with the convergence of lines extended in depth, which converge on the horizon. The most prominent convergence point, the *pvp*, is a strong center for the composition, drawing the gaze into the depth of the scene geometry.

4.2.1 Definitions

To assure that the tiled floor is a simple, strong, depth structure, we observe the following constraints.

1. The view direction is perpendicular to the image plane.
2. The physical tile boundaries are perpendicular and parallel to the image plane.

Prior to describing how to draw a tiled floor, several terms (Figure 4.3) must be defined.

The **baseline** lies along the bottom of the canvas frame, meeting the ground plane. Intersecting the baseline, perpendicular to the image plane, is the physical ground plane on which the viewer/painter stands (Figure 4.2).

The **horizon** is above and parallel to the baseline at the viewer's eye level.

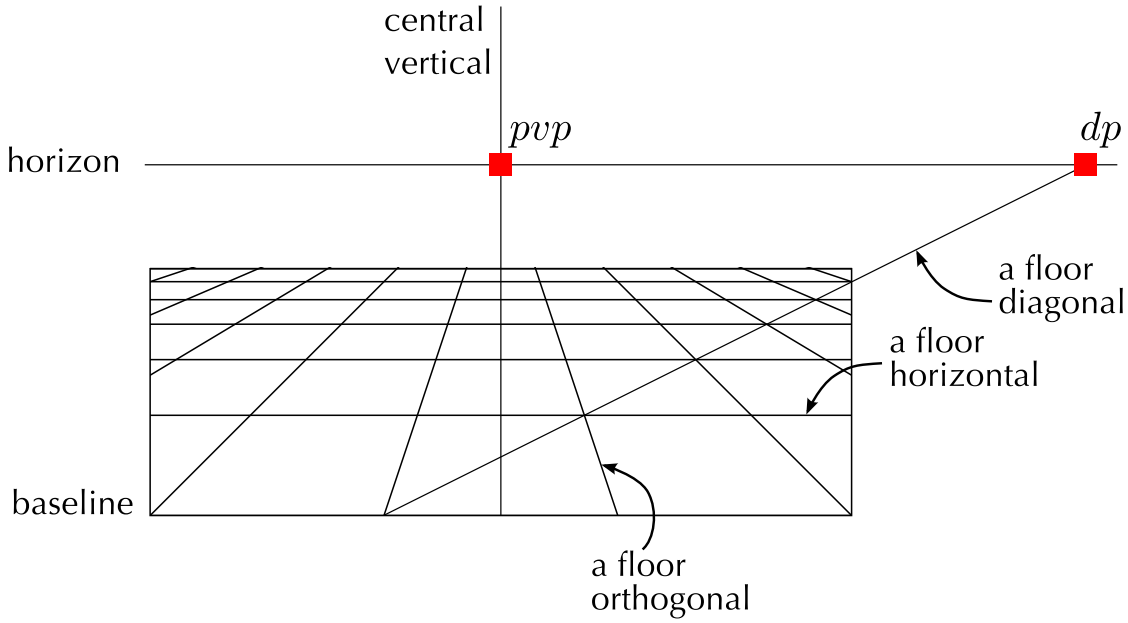


Figure 4.3: The tiled floor. The terms used to construct it in perspective are labelled.

The **principal vanishing point**, *pvp*, is the point on the horizon in the direction of the viewer's gaze. Alberti calls it the *centric point* [37, p. 25].

The **distance point**, *dp*, is on the horizon. The distance between the *pvp* and the *dp* is the distance of the viewer from the canvas.

The **floor orthogonals** are perpendicular to the image plane in the physical space. In the perspective space they converge on the *pvp*.

The **floor horizontals**, or transversals, are parallel to the image plane in the physical space. In the perspective space they remain parallel while the distance between them foreshortens with increasing scene depth.

The **distance point diagonals**, or **floor diagonals**, are lines from the *dp* to the baseline. They pass through tiles corners, determining the foreshortening of the floor horizontals.

The **central vertical**[2] is the perpendicular line from the baseline to the *pvp*.

The *pvp* and *dp* fix the viewpoint with respect to the canvas. The floor orthogonals, horizontals and diagonals are determined by the *pvp* and *dp* and the size of the tiles.

[2]Although not required for the tiled floor, the term is useful in describing the perspective geometry of a picture.

4.2.2 Distance Point Construction

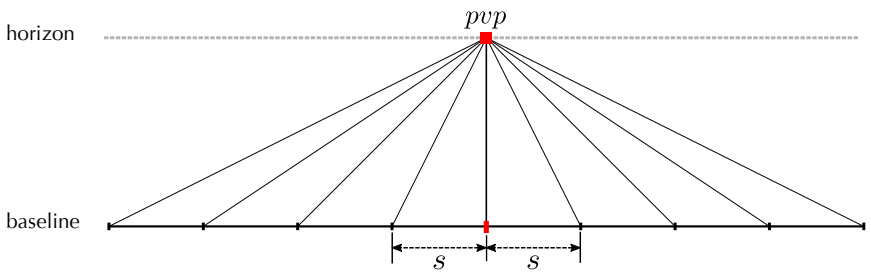


Figure 4.4: The orthogonals connect tile corners on the baseline to the *pvp*. The distance between the floor orthogonals at the baseline is the physical size of each tile.

On the canvas the tiled floor is the perspective grid of the 3D scene. The physical tiled floor consists of squares of a fixed size, s. Because the horizon is at the viewer's eye level, the line from the *pvp* to the viewer's eye is perpendicular to the canvas. The perspective tiled floor lies inside the triangle formed by the baseline and the *pvp*. *The distance point construction* that implements the transfer of a tiled floor physical 3D geometry to the 2D canvas is now described.

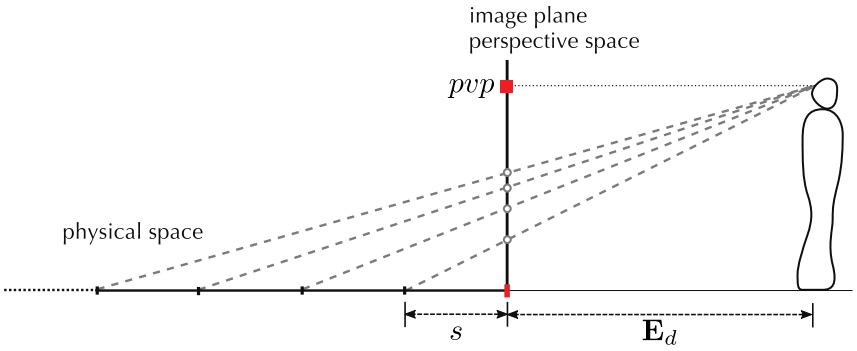


Figure 4.5: Side view of the 3D view geometry between the physical tiled floor, the image plane and the viewer: placing the floor horizontals on the canvas is trivial.

To make the floor, place tile corners on the baseline at regular intervals, s, starting with the middle one on the central vertical[3], as illustrated in Figure 4.4. Draw lines from them to the

[3]For a *pvp* vertically centered inside the frame, this constraint only depicts floors with an even number of tiles on the baseline. A slight variation of the construction (Appendix E.2) permits depicting tiled floors with either an even or odd numbers of tiles on the baseline.

pvp marking where the infinite-length floor orthogonals project onto the canvas. To complete the perspective tiled floor, floor horizontals must be drawn, the number of which depends on the depth of the scene. In the assumed scene geometry, they are parallel to the baseline. To determine their placement on the canvas, the distance between the viewer and the picture plane, $\mathbf{E}_d$, matters.

Given $\mathbf{E}_d$, the positions of the floor horizontals are easily evaluated. Figure 4.5 illustrates the side view of the 3D view geometry: the viewer stands on the ground plane $\mathbf{E}_d$ from the image plane. In the side view, each distance marked behind the image plane is a depth division of the physical tiled floor. The rays to each from the viewer's eye intersect the image plane at different heights: on the image perspective space, those heights are the positions of the floor horizontals in the scene depth.

While the construction is geometrically clear and easy in the 3D side view, the actual construction, Figure 4.4, is different. The viewer must be in the canvas perspective space, along with the depth divisions of the physical tiled floor. The distance point construction includes the 3D side representation in the image plane perspective space, a 2D geometrical solution. Renaissance artists envisaged Figure 4.5 superimposed on Figure 4.4 by a 90° rotation (Figure 4.6). Figure 4.7 shows the result, which is the essence of the distance point construction.

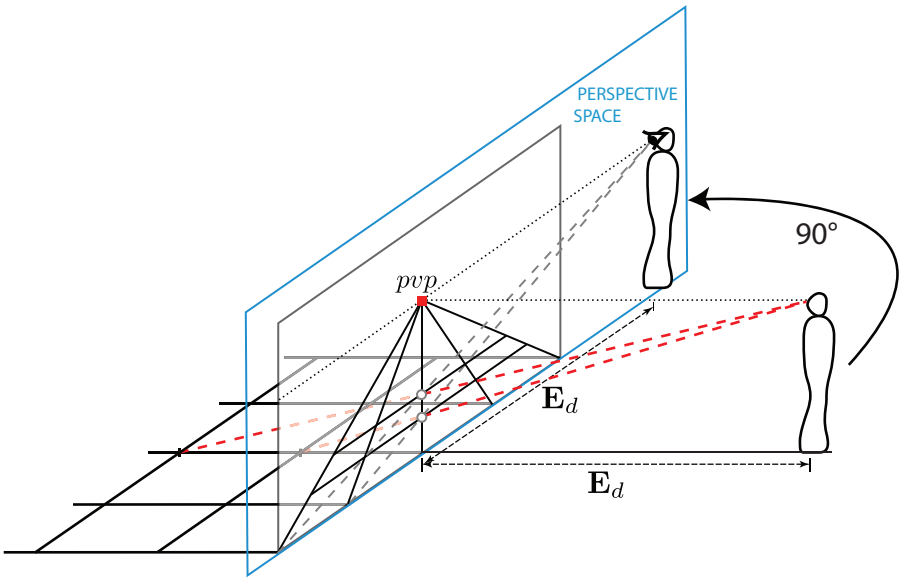


Figure 4.6: The physical plane geometry that determines the heights of the floor horizontals on the canvas (Figure 4.5) is rotated by 90° around the central vertical to bring the physical space geometry into the canvas perspective space.

In practice, the construction is accomplished as shown in Figure 4.7. On the canvas the distance point, dp, representing the eye, is placed on the horizon, at a distance from the pvp that is the distance between the viewer and the image plane. Then the floor orthogonal that is on the central vertical functions as the image plane of Figure 4.5. This particular floor orthogonal is called the *image plane reference*. It splits the canvas in two: one side containing the dp, the other containing half the baseline tile corners. The superimposition depends on the tiles being square, so that baseline measures match physical measures. From the dp, rays are drawn through the image plane reference to each baseline mark. The intersection heights are the projection of a tiled floor depths onto the image. The floor horizontals are parallel to the baseline, passing through the intersection heights. The rays from the dp to the baseline are also the floor diagonals.

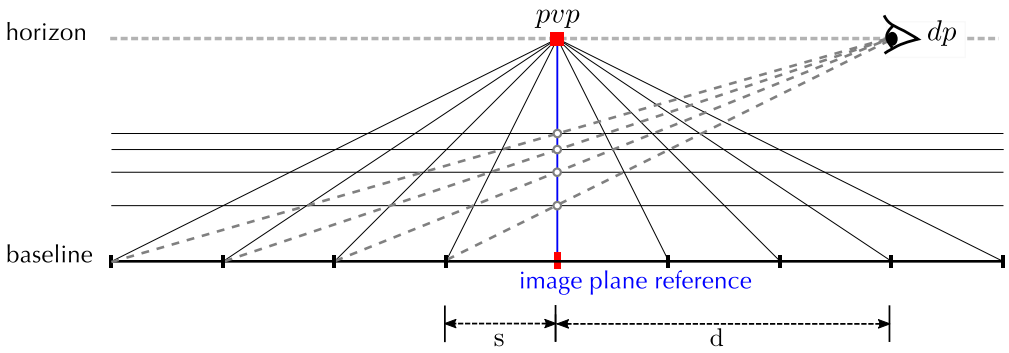


Figure 4.7: To construct the floor horizontals, the distance point, dp, is located on the horizon. On the image plane reference, the intersections of the rays from the dp and to the baseline marks determine the heights of the perspective floor horizontals.

This completes the floor construction. The key elements—the horizon, the baseline, the pvp and dp—are established, encoding the viewer's relation to the scene and canvas. From them the construction builds the tiled floor, which provides a strong perspective grid on the canvas, on which the further development of the picture is based.

The distance point construction requires a specific scene geometry. In addition to the two constraints applied to the 3D viewpoint (Section 4.2.1), there must be a floor orthogonal on the central vertical. As shown next this later requirement is unnecessary, but facilitates the visualization of the concordance between the 3D scene geometry and the practical one used on the canvas.

4.2.3 Piero's Diagonal Construction

This section presents a different tiled floor construction based on the distance point geometry: *the diagonal construction* of Piero della Francesca. This alternative demonstrates that using the distance point geometry to construct a tiled floor is robust and flexible.

Piero's diagonal construction (Figure 4.8) has the benefit of succinctness. Piero realized that a single floor diagonal can construct the perspective floor because every diagonal passes successively through square tile corners. He selected a floor diagonal connecting the dp to the furthest tile corner on the baseline. This choice is effective because it gives the depth locations of all floor horizontals for a square floor.

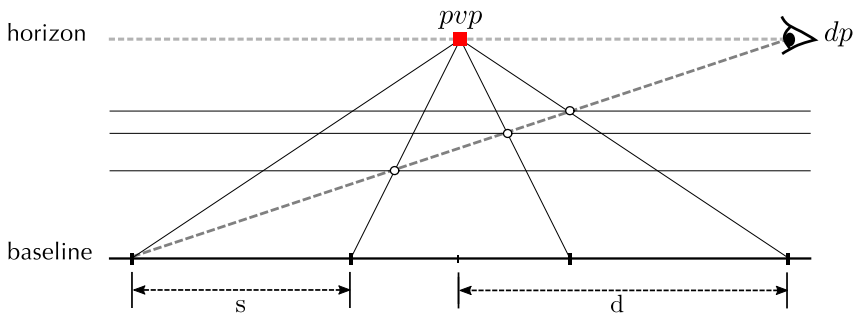


Figure 4.8: Alternate construction of a tiled floor using Piero's diagonal construction. A single floor diagonal runs across the entire picture: successive intersections with floor orthogonals give the floor horizontal locations on the canvas. Tile edges aligned with the central vertical are not needed.

Furthermore, Piero's surviving diagram for his diagonal construction (Figure 4.9) is quite remarkable. It shows the perspective view of the tiled floor with the reflected physical floor immediately below: the two spaces involved in the practice of constructing a realistic image. His juxtaposition demonstrates the correspondence between the perspective and physical floor diagonals.

In effect, Figure 4.9 emphasizes the convergence of the perspective floor orthogonals, which are parallel in the physical space, in contrast to the floor horizontals, which are parallel in both spaces. All other floor lines that are parallel in physical space converge somewhere on the horizon. In addition, Piero's construction makes prominent the common baseline, the line that is invariant under perspective projection, and therefore useful for transferring measures.

Piero's diagram demonstrates that Renaissance artists had a deep understanding of perspective projection as a practical solution of simulating depth for a tiled floor. The nature of

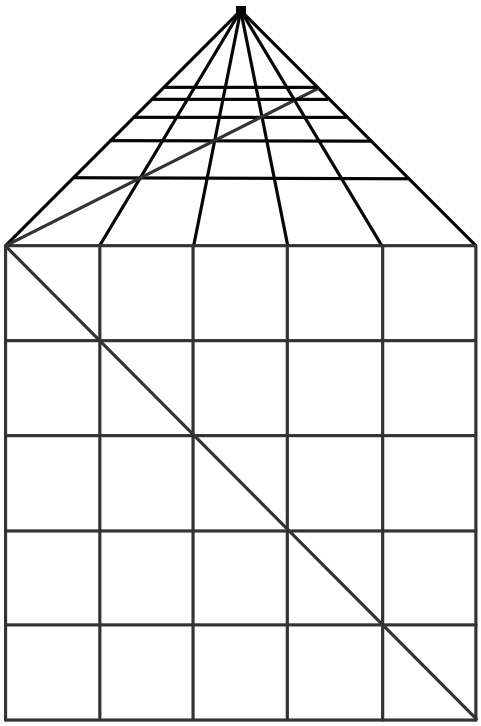


Figure 4.9: Reproduction of Piero's diagram of his diagonal construction to draw a square floor of square tiles in perspective. The diagram juxtaposes the image of the tiled floor with the physical floor represented, reflected, underneath. The floor diagonal running across the tiled floor extremes is shown in both parts: it evaluates the locations of the perspective floor horizontals; it highlights the two floors geometrical connection.

the lines as they are and as they are seen are differentiated: the revision of Euclid's parallel lines is made clear, parallel lines can intersect in the perspective space of a realistic image (c.f. Section 1.2.2).

Appendix E presents two more variants for constructing a tiled floor in perspective. The diversity of methods created during the Renaissance demonstrates artists' intuitive understanding of projective geometry. The relative benefits of each method are described in Appendix E.4. While being geometrically equivalent each has different practical properties for manual execution of the tiled floor geometry.

4.2.4 Applications to the Design of the Framework

The tiled floor constructions described in Sections 4.2.2 and 4.2.3 encompass concepts that were essential for the development of my framework.

In both constructions the baseline has a critical role: it is a line, invariant under the perspective transformation, where the world of the picture and the scene to which it corresponds coincide. Thus, it is ideal for transferring measurements. In the distance point construction, the 3D ray intersections are reproduced on the canvas using the baseline. In Piero's diagonal point diagram, both the floor orthogonals and the floor diagonal of the perspective picture meet their physical equivalents at the baseline.

Most importantly, the tiled floor constructions reveal two folds used to simulate realistically the geometry of a 3D scene on the canvas. With the tiled floor constructions, Renaissance artists and Piero in particular, discovered folds that map important half-planes of the 3D space onto the image plane. Figure 4.10 illustrates these two foldings, which are half-planes rotations by 90° about the central vertical and the baseline: they decompose the 3D geometry of projection into geometric evaluations executed directly on the canvas. The 2D computations of my framework (Chapters 5 and 6) are based on them.

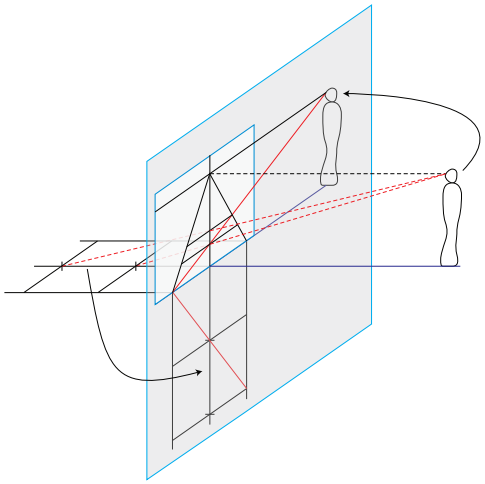


Figure 4.10: The distance point construction folds the side view of the 3D geometry around the painting to the image plane: the half-plane with the viewer's eye, feet, line of sight, and physical tiled depth divisions is rotated by 90° about the central vertical (upper-right arrow). Piero's diagonal construction folds the 3D ground plane behind the painting to the image plane: the half-plane with the physical tiled floor is rotated 90° about the baseline (lower-left arrow).

In particular, Piero's fold is fundamental to my thesis. The juxtaposition of the physical floor below the perspective one reinforces the coexistence of the 2D and 3D spaces, how they are entangled with and complement each other. As shown in the remainder of this thesis, Piero's placement in the image plane of the perspective floor and of the physical floor is the geometric structure needed to simulate perspective differently, that is on the flat canvas surface. Among others, it enables

- mapping any ground plane point from the physical floor to its perspective representation (Chapter 5) and
- deriving algebraic transformations that map canvas images to other images under rotation, translation or scaling occurring in the perspective space (Chapter 6).

Finally, because my framework 2D computations are based on these rotations of the 3D space into the canvas they match the computations normally performed in the 3D geometry, as shown in Appendix B.

From a practical point of view, my framework computes geometrically the tiled floor according to a close variant of the distance point construction described in Section 4.2.2. Appendix E.2 presents the construction my framework uses: it is geometrically equivalent to the distance point construction, and does not required the central vertical constraint[4]. In addition Piero's construction with its single diagonal, is a simple visualization in my framework providing a concise sanity check that the floor displayed is correct.

The rest of this chapter presents a few artists' construction line examples that depict architectural volumes and projected shadows.

4.3 Scene Elements

The geometry of the perspective tiled floor is the base of the composition constructed on the canvas. While the construction of a floor made of square tiles extends to rectangular, non-rectangular and patterned tiled floors, its basic purpose is not to impose an appearance on the ground plane. The lattice of the tiled floor, the repetition pattern of its tiles, is above all a reference structure for organizing the depth volumes of the scene, especially the architectural forms that define the bounds of the composition volume.

[4]The different aspect of its construction naturally follows from the distance point construction presented in this chapter, but its construction is initially less visually intuitive in term of 3D geometry.

This section discusses the potential to pattern the tiled floor and presents selected diagrams of Piero della Francesca that illustrate the construction lines used for complex perspective volumes. Those examples demonstrate that the 3D simulation of the picture emerges from the construction lines on the ground plane: base geometries on it are elevated above it. The constructions are not described; this section gives only their flavour, as a proof of concept showing that the lines on the ground plane is the basic construction tool from which Renaissance artists developed complex scene volumes while working directly on the canvas. Chapter 5 describes how such line construction methods translate into algorithms for my framework.

Floor Patterns. A regular tiled floor can be subdivided into complex floor patterns. Connecting depth intersections can create interesting and decorative floors. Artists included such floors in their images, relying on dense triangulations of the tiled floor depth grid. Figure 4.11 shows an early example of a complex pattern on a tiled floor (Ambrogio Lorenzetti *The Presentation in the Temple*, 1342). The reproduction was reconstructed by Richard Talbot, who noticed that floor patterns in perspective appear at least a century before paintings with fully developed perspective [103]. Those observations support Talbot's convincing claim that artists discovered how to depict realistic perspective images on the canvas through "playing with" geometric patterns on tiled floors [103].

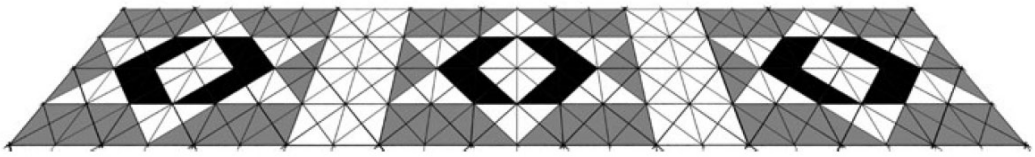
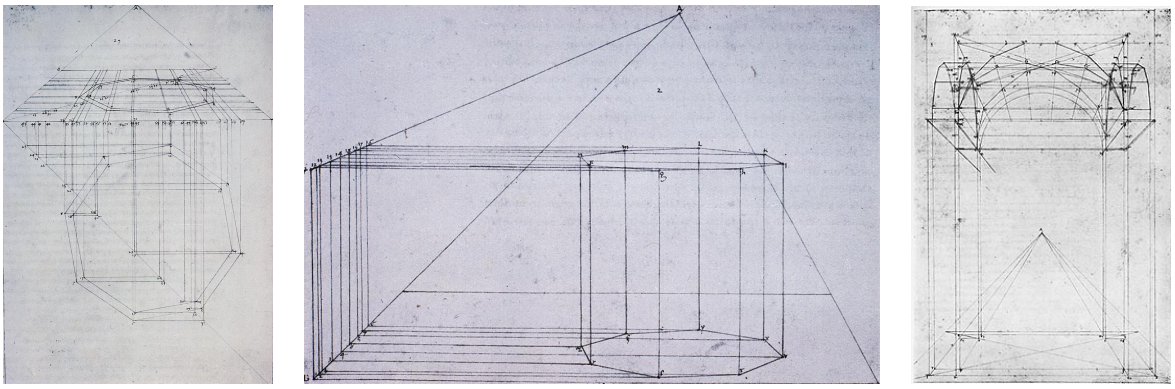


Figure 4.11: Richard Talbot's reproduction of the complex tiled floor pattern that appeared in Ambrogio Lorenzetti's painting, *The Presentation in the Temple* (c. 1342) [103].

Indeed, Piero's book, *De Prospectiva pingendi*, 1482, contains complex diagrams showing constructions to elevate volumes in perspective, three of which appear in Figure 4.12: (b) shows how to build an octagonal structure starting from its geometry on the ground plane (a); (c) shows the construction of a vaulted bay. They demonstrate the power of the distance point geometry and of the correspondence between the two floor spaces on the image plane.

Octagonal Structure. Figure 4.12 (a) shows the construction to draw in perspective an octagonal base from its physical floor representation. Arbitrarily placed on the physical floor



(a) (b) (c)

Figure 4.12: Piero's perspective volumes illustrations: the first two to draw an octagonal volume; the last one to draw a vaulted bay.

it is slightly rotated such that no edge is parallel to a tile edge. Elementary constructions of Euclidean geometry transfer the vertices to the perspective floor using the invariance of ratios under projective transformation: the measures are taken along the diagonal running across the physical and perspective floors.

Figure 4.12 (b) then shows the perspective construction of an octagonal prism. The boundary of the perspective tiled floor is drawn. The floor orthogonal from the *pvp* to the farthest bottom picture corner is an important location for the elevation of the prism: the depth of the base vertices are transferred to it; the foreshortened heights are evaluated so to be transferred back above the base vertices.

Vaulted Bay. Figure 4.12 (c) shows the construction lines to depict a vaulted bay in perspective. This bay closely resembling the main architecture component of Masaccio's painting, *The Trinity*, (c. 1426). The depiction in perspective of the arch uses an elaborate set of construction lines to approximate the foreshortening of the curved bay. In particular, the smooth arc at the back of the bay, which defines the curvature of the central arch, is mapped piece-wise to the front of the bay, using construction lines on the ceiling. Indeed, the ceiling contains prominent orthogonal lines in the manner of a tiled floor. The arches on either side of the nave stand on volumes projecting from the wall, have also their lateral foreshortening evaluated. The diagram is too complex to describe. But it shows some floor construction lines on the floor, many more on the ceiling plane and just below it, as well as on the vertical planes spanning them at the edges of the floor.

In these diagrams, Piero demonstrates the importance of construction line intersections with the floor boundary, the baseline, the most lateral floor orthogonal, and the diagonal running across the width of the floor. In particular, it is important to use the floor orthogonal farthest from the center. Doing so avoids degeneracies and optimizes consistency and accuracy. Always measuring from the same small set of construction lines minimizes compounding error. Using the longest orthogonal and intersections close to 90° are both good for increasing precision. Piero's diagrams isolate the axes that are the special locations where measures are transferred between the planes of the two spaces—the physical and perspective tiled floors—and between the various planes of the perspective space itself.

The construction methods shown in these three diagrams demonstrate that practical methods for constructing projections of complex structures were known in the Renaissance. They rely on the tiled floor, on the geometry of which all architectural constructions depend. The result is a coherent perspective simulation of a 3D scene created solely from 2D constructions. Chapters 5 and 6 describe the relevant geometry concretely and its algebraic expression, which is suitable for computation.

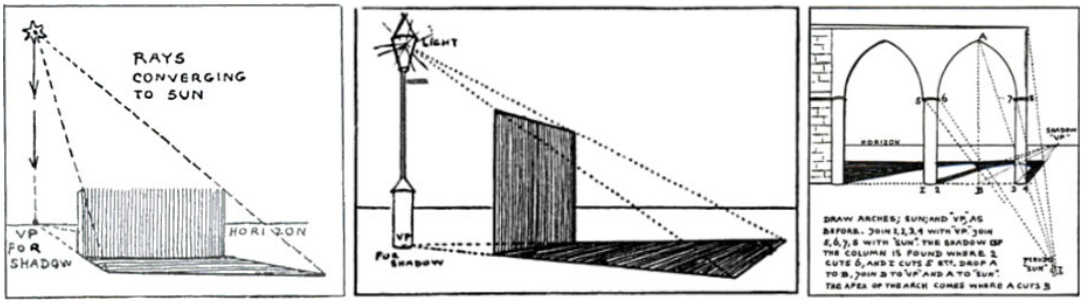
4.4 Projected Shadows

When constructing volumes in the perspective picture, the geometric connection with the appearance of shadows became apparent. Some shadows can be modelled with canvas construction lines. In effect, Figures 4.13 (a) and (b) illustrate line constructions to create projected shadows of a simple object, a vertical wall. Two sets of rays are involved:

1. the ground rays, which join on the ground the light to the object corners and which define the direction of the shadow; and
2. the light rays, which join off the ground the light to the object corners and which define the size of the shadow—its depth limit.

Both sets of rays are extended beyond the object to where they intersect. Since the wall is vertical and touching the ground, the two intersections produced by the light and ground ray pairs for each of the two vertical edges, suffice to define the shadow.

Figure 4.13 (c) illustrates a shadow construction for a complex geometry. An architectural structure made of pillars supporting two arches creates a shadow towards the horizon due to an illumination in the back of the artist.



(a) (b) (c)

Figure 4.13: Cole's examples of line constructions that produce projected shadows.

The examples of Figures 4.13 are taken from Cole's manual [17], written in the early twentieth century. The manual compiles many artists' constructions for drawing objects in perspective and contains a full chapter on the 'Perspective of Shadows' that describes the geometry of projected shadows. Cole's diagrams and descriptions are informal and incomplete: he expects the artist to complement them with the tacit knowledge of a skilled practitioner. Chapter 5 describes the construction line algorithms incorporated in my framework, which implement the constructive geometry of shadows. Essential for creating projected shadows is to understand how different illuminations are placed on the picture, which is described next. Lastly, the interests of projected shadows in the Renaissance and in my framework contexts are discussed.

4.4.1 Illumination

Artists construct projected shadows from illumination references on the canvas. They may not have in mind an exact 3D location for the light, but they know where to place it on the canvas to create the shadows they want for the composition. Artists position the light source and the center of projection of the ground rays, called the light vanishing point: the two points encode the minimum information needed to simulate 3D in this double projection problem.

These two points have different relations on the picture according to the type of illumination lighting the scene. Figure 4.14 shows possible illumination positions relative to the scene and the artist with his canvas. The source of illumination is either behind or in front of the artist.

For example, illumination in front of the artist may have a location on the image plane, or may be off to the side. In the first instance the artist places the light source on the canvas in accord with the composition, the light vanishing point directly below: a distant light, such as the sun, has its vanishing point on the horizon (Figure 4.13 (a)); a nearby light has it within the area of the ground plane (Figure 4.13 (b)). Lights illuminating a scene from the side or from high overhead are rarely within the picture frame: a directional representation works best.

Most interesting is a source of illumination in the back of the artists, behind them (Figure 4.14). Artists discovered that a suitably placed pseudo-source casts appropriate shadows. Indeed, the source of illumination projects onto the image plane with the viewer's eye's as the center of projection: a sun on the left side behind the artist projects on the canvas to the right of the *pvp*, as illustrated in Figure 4.13 (c). Sources of pseudo-illumination also have vanishing points, which are the projection through the viewpoint on the canvas of their anchor point on the ground plane. Because projection reverses up and down they fall on the canvas above the corresponding pseudo-light.

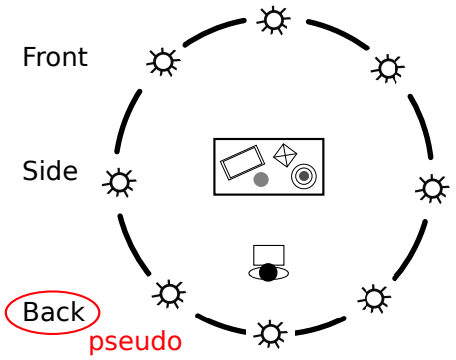


Figure 4.14: Illumination positions with respect to the placement of an artist drawing a simple scene. Redrawn from an illustration in a modern art manual [7, p. 60].

Given the illumination on the canvas, the construction lines that create shadows resemble those used to simulate 3D in the perspective picture. My framework, described in the next chapter, describes the geometric construction for projected shadows.

4.4.2 Contexts

Only *projected shadows*, also called perspectival shadows [4], which are instances of an object projection on a flat surface, are relevant to this thesis, as they can be geometrically depicted[5]. There are two reasons why shadows are part of this thesis: they were of geometric interest to Renaissance artists and they are an interesting double projection problem for my framework.

[5]Shading of surfaces due to differential illumination and self-shadowing are not considered as artists rarely use construction lines for these effects.

Shadows puzzled Renaissance artists who carefully observed them to understand the properties of projection and the nature of light [64]. Before the camera, projected shadows were a dramatic example of 3D objects projected to a 2D surface. As in antiquity, Renaissance artists-mathematicians experimented with the flat shape of the shadow created by a source of illumination, which they examined from the viewpoint of the light. The variety of shadow effects, varying with the light type and environment conditions intrigued Leonardo da Vinci: the smoke-like boundaries of shadows challenged for him either the closure of the object or the straightness of light rays. Albrecht Dürer discovered it was possible to project shadows using perspective constructions. However Renaissance artists drew shadow effects lightly, so as not to distract the viewer's attention from the main content. Shadow geometric constructions, which were only explicitly described in the nineteen century, use concepts familiar from the perspective constructions, but there are some additional difficulties. The mathematician Taylor (1685–1731) described the construction of shadows in the picture plane as 'putting the . . . Rules of *Perspective* into Perspective' [4, p. 529].

Indeed, for my framework depicting shadows is interesting because it requires double projection: projecting an object to a surface by a source of illumination, projecting the patterned surface to the image plane. In addition, projected shadows anchor objects in depth: they provide the viewer with a strong perception of the scene geometry. To an image observer an object's location in 3D scene is more precise if it has a shadow falling on the ground [14]. While the position of a floating sphere is ambiguous, its projected shadow onto the ground resolves the ambiguity. Finally, shadows create compelling interactions between 2D and 3D objects, and thus are essential for creating compositions that captures the 3D/2D dynamism among the depicted scene elements.

Chapter 5

The Framework

The thesis framework is the computer program I developed to support both image composition and realistic depth depiction. This chapter presents the framework by describing in turn its components—planar pattern, tiled floor, objects and projected shadows—and how they interact with one another. The framework provides a tiled floor construction, which specifies a depth grid for the 3D scene. The tiled floor is constrained within the 2D geometrical pattern overlaid on the image plane. The constraints produce a 3D ground plane pattern, which also has a 2D pattern, creating a simulation of depth with planar relationships. Within the overlaid patterns, the planar pattern and the perspective tiled floor, is specified the picture content: perspective volumes, planar objects, and projected shadows, which are depicted and arranged to create a perspective picture composed on the plane.

The user positions the volumes, planar objects, and the light; the framework renders the scene, simulating 3D. The framework constructs the perspective picture by evaluating line intersections in the picture to simulate the scene depth geometrically. This chapter describes a geometric approach that derives the perspective picture from construction line algorithms. The algorithms automate the artists methods presented in Chapter 4 along with new constructions that rotate a panel and to construct shadows on vertical walls. These results are the practical geometrical contributions of the thesis and have been published [38, 39].

Chapter 6 duplicates algebraically the geometrical calculations of this chapter: an alternative solution that uses 2D projective transformations to implement 3D depiction on the canvas, simplifies the framework code. The formalism presented there is a theoretical contribution of the thesis.

5.1 Overview

This chapter describes my framework and explains the constructive geometry algorithms that evaluate the schematic representation. In my framework application, a user creates a picture by manipulating framework components, which I first introduce in this section as a whole before turning to the details of each in the rest of this chapter. The framework components are:

- a pattern, that provides focal points for planar composition;
- a tiled floor, within the planar pattern, for simulating 3D relationships on the canvas;
- objects populating the scene, of one of the two following forms:
  - simple geometric solids that fill volumes occupied by architectural features; and
  - panels that are semi-transparent images approximating local foreground objects;
- a source of illumination and the shadows it produces.

Figure 5.1 displays the interface of my application, showing at the center the canvas where the schematic representation is composed. The canvas contains an early design of a picture. A planar pattern is first selected to provide the 2D structure for the composition. The planar pattern organizes the 2D Euclidean space of the canvas, providing 2D locations where scene objects can be placed to form simple geometric figures, controlling alignment, visual balance and rhythm. Then, the tiled floor is integrated within the planar pattern, providing a structure that sustains scene depth. Constraints on the placement of floor elements provide visual harmony between scene depth and planar composition.

Once the depth structure of the floor is set within the planar pattern, the composition space is populated with scene objects. The intermingling of planar and depth structures assures the dual nature of each scene object, its physical 3D reality and its role in the 2D composition. There are two kinds of objects: architectural volumes and foreground figures. The distinction follows the different treatments of structural shapes and organic forms by Renaissance artists (Chapter 3). The user places editable representations of scene objects at canvas landmarks. The framework uses 2D construction lines to update their appearance. Architectural volumes are geometrically constructed by elevating floor tiles. Foreground figures are the content of panels, externally created images made available on the left side of the canvas (Figure 5.1).

Finally, projected shadows cast by, and falling on, objects in the composition can be added. The user positions the sources of illumination on the canvas, and then line construction methods depict projected shadows. Adding shadows creates valuable perceptual interactions between the planar and volumetric objects of the composition.

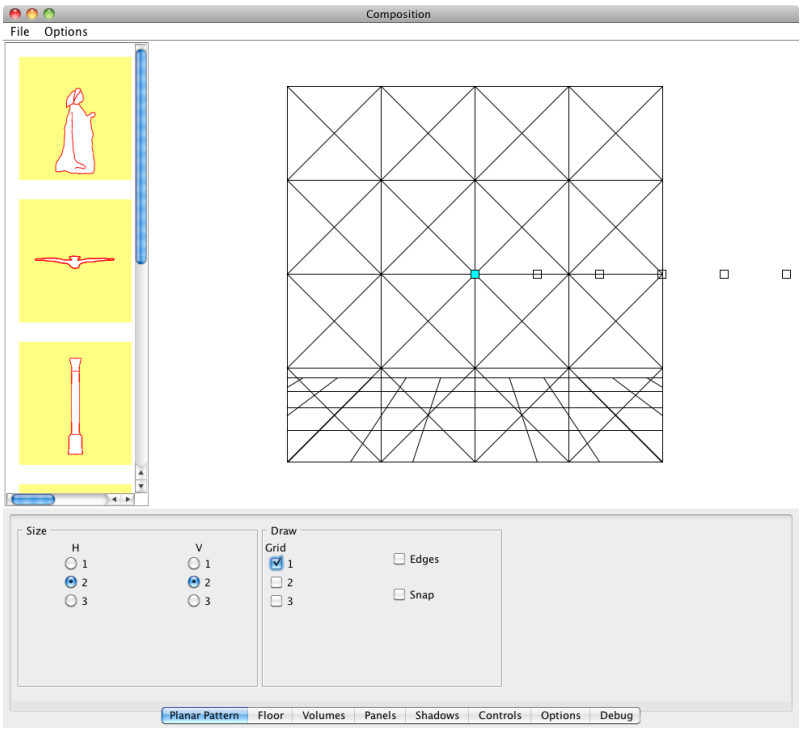


Figure 5.1: The graphical interface of my framework application. The left column contains images to be used as panel content. The central region is the canvas, where the schematic representation is created. Below, tab-panes contain specific controls for each of the framework components.

Complementing this chapter, implementation details including the user interface description are in Appendix D. Rendering is not discussed in this chapter: Chapter 7 describes the painting algorithm discovered in my research.

5.2 Intermingling the Planar and Depth Structures

Artist and art historian Richard Talbot demonstrated the excellent fit of a specific planar pattern the 2D geometry aligns of which with the composition of acknowledged masterpieces of the Renaissance [102]. His planar pattern naturally includes the Renaissance tiled floor perspective construction within it. Thus, in my framework Talbot's pattern is used to ensure a focus on planar composition, providing an underlying planar structure that guides composition and to construct a tiled floor whose perspective appearance integrates elegantly within.

This section describes first Talbot's pattern, its geometry permitting the creation of relationships between salient features of a picture. Other patterns than Talbot's could easily be used. Then, the constrained construction of a tiled floor within it is explained. This framework management enables us to visually connect planar composition and realistic depth as strategic locations on the tiled floor and the depth ratios blend with the planar geometry of the picture.

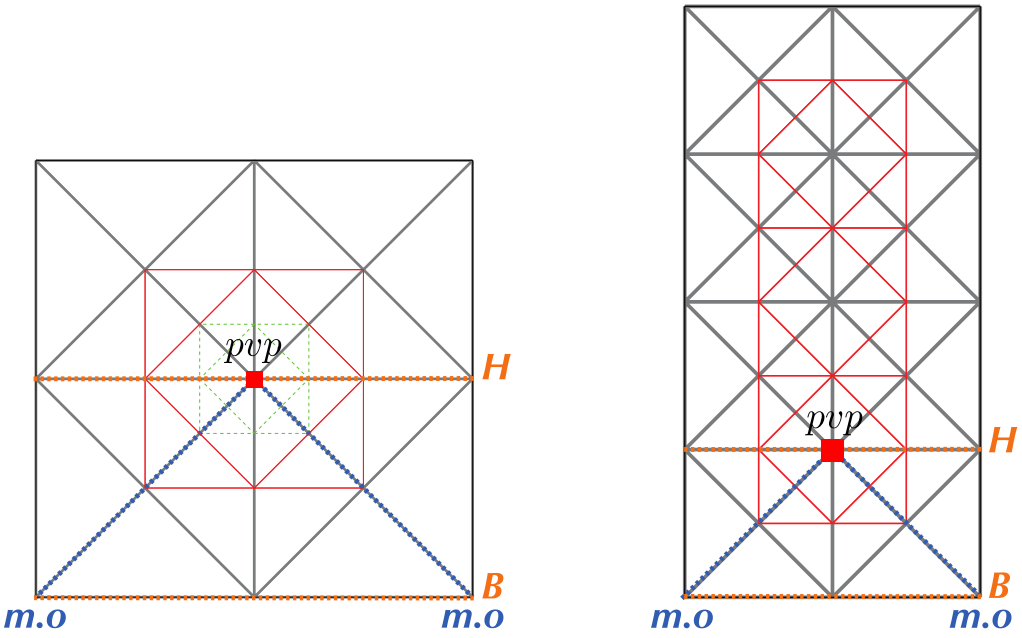
5.2.1 Planar Structure: Talbot's Picture Pattern

My framework uses Talbot's planar pattern rather than the usual rectilinear grid of imaging software. Talbot's basic pattern is shown in Figure 5.2 (a). At the largest granularity (gray) is the main pattern: a square, with its four fold reflection lines, plus a diamond defined by the reflection lines inscribed in the main square. The pattern is recursive. The second level (red) is the same geometry but smaller, inscribed in the main diamond of the first level. The recursion applies further (green at the third level). The squares and diamonds inscribed in each other share a single set of reflection lines.

The reflection lines of the squares are the vertical, horizontal and diagonal bisectors of the square. The horizontal and vertical lines intersect the centers of the edges of the inscribed squares, while the diagonal pair connects opposing vertices. The vertical and horizontal reflection lines in turn pass through the vertices of the inscribed diamonds, while the diagonal lines pass through the centers of their edges.

The basic pattern may be multiplied to create larger and more complex compositional structures, such as vertical stacking of two basic patterns which is a layout suitable for depicting the nave of a church (Figure 5.2 (b)). Inscribed squares may be inserted between basic patterns to unify otherwise separated portions of the 2D space. In Figure 5.2 (b) it is possible to see the pattern as two stacked basic patterns or three intertwined ones. The combination of squares and diamonds emphasizes two types of intersections: the corners of the square and the diamond vertices at the mid-point of square edges. Thus at each level, eight regular locations are distinctively marked on the boundary. Inscribed shapes reproduce those diverse intersections at denser levels, recursively nested inside. Ratios of the sides of inscribed squares are $1:2:4:8\ldots$ creating a geometric sequence, as do the sides of inscribed diamonds.

Studying the composition of several Renaissance paintings, Talbot found the intersections of his pattern to be locations of dominant figures and of architecture lines. His planar pattern contains spatial landmarks that are visually salient. For example, selecting three points to be the vertices of an isosceles triangle is easy. Indeed, many of the typical Euclidean shapes found



(a) (b)

Figure 5.2: Talbot's planar pattern for image composition. In (a) the basic pattern is shown. In (b) a stack of two provides for a vertically oriented frame. On both, the recursively inscribed patterns within the main level are visible, with three levels in (a) and two in (b). The labels indicate the default positions of the tiled floor key elements.

in the composition of Renaissance paintings—circles, triangles or even five pointed stars—occur in Talbot's patterns. Another example is the correspondence for the main volume of an enclosed room, where ceiling lines for the side walls can use planar pattern intersections, such as the horizontal vertices of second level inscribed squares. (A wall feature is described in Appendix D, § D.2.3.)

5.2.2 Depth Structure: Perspective Tiled Floor

We saw in Chapter 4 that Renaissance artists required a simple depth structure to simulate a 3D scene directly on the canvas. They used a tiled floor consisting of square tiles aligned perpendicular to the image plane and parallel to the baseline. The framework user specifies the configuration for such a tiled floor: the locations of the *pvp* and *dp*, which set the principal perspective governing the geometrical appearance on the canvas, and the discretization, the

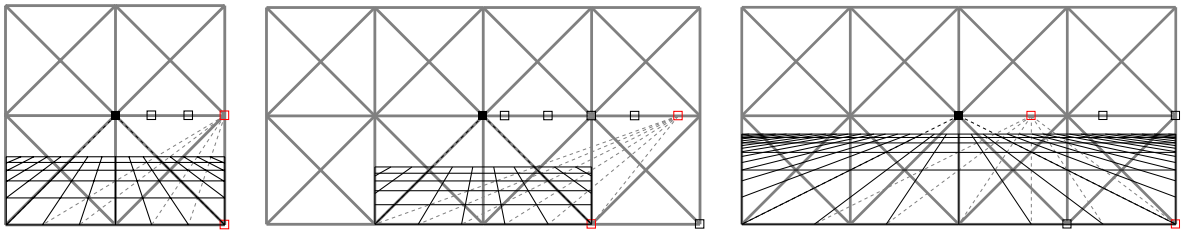
numbers of tiles in the baseline and in depth, which define the granularity of volumes in the schematic representation.

My framework restricts the potential placements for the tiled floor elements to focal locations of Talbot's planar pattern, so that the arrangement in depth reinforces the image plane composition. The placement constraints are described next.

Constraints. The important elements of the tiled floor, defined in Section 4.2.1, are revisited and placed at their most common location in Talbot's pattern. (The labels used those of Figure 5.2).

- **The horizon**, H, which sets the viewer's eye level, coincides with the horizontal reflection line of the bottom row of squares.
- **The principal vanishing point**, *pvp*, which sets the viewer's line of sight, is at the horizontal center of the planar pattern.
- **The baseline**, B, which connects the perspective space with the physical one, is at the lowest horizontal line of the planar pattern.
- **The main orthogonals**, m.o, which are the perspective tiled floor's most oblique orthogonals, intersect at the baseline with vertical lines of main square(s).
  The main orthogonal planes are the vertical planes that contain the floor main orthognals.
- **The floor orthogonals**, lying between the main orthogonals, reach the baseline at regular intervals that depend on the tile size.
- **The distance point**, *dp*, which is separated from the *pvp* by the viewer's distance from the image plane, is on the horizon at a location vertically aligned from a baseline tile corner.
- **The floor horizontals**, which show depth foreshortening and are situated below the horizon are within the lowest half of the bottom row of the main square(s).

This list adds the main orthogonals and their planes to the list in Section 4.2.1. While Chapter 4 emphasized the ground and image planes, another plane orientation is also important: the plane orthogonal to both the ground and image planes is needed for height constructions (§ 5.4.1). Since the intersections of those planes with the ground plane are floor orthogonals, they are called the *orthogonal planes*. The useful ones are those most oblique in perspective, which reach the frame corners. Piero used them in an important construction that elevates a volume above the tiled floor, Figure 4.12.



(a) (b) (c)

Figure 5.3: Three perspective tiled floors within Talbot's pattern: the *pvp* is at the center; potential *dp* and picture frame right corner locations are the small outlined squares on the horizon and the baseline respectively. (Selected ones are in red.)

The constraints on the key elements ensure that the horizon line, the *pvp*, the baseline, the floor main orthogonals and the *dp* are all placed at important locations in Talbot's planar pattern. The placements provide a strong center, the *pvp* at the horizontal centre of the picture, which is also the vertical center when the pattern is not extended in height. *A posteriori*, one can deviate from this powerful use of the center, but it is advantageous to over constrain initially, in order to provide maximal unity within the planar composition[1]. The perspective tiled floor creates a geometric rhythm, a strong visual cue for recession in depth. Constraining the tiled floor within Talbot's planar pattern merges geometric visual rhythms in depth with linear and geometric ones in the frontal plane.

Examples. Subject to the constraints the user chooses a tiled floor configuration. The framework displays the potential placements from which the user can choose (Figure 5.3). The framework draws the tiled floor in perspective by computing the line intersections using a distance point construction (§ E.2). The infinite floor orthogonals of Figure 4.7 are clipped to the deepest floor horizontal, producing tiled floors covering only part of the region between the baseline and the horizon (Figure 5.3).

As is common in early Renaissance art, the distance point is by default on the vertical edge of the picture frame (Figure 5.3 (a)). The *dp* can be anywhere on the horizon, but the constraints ensure that its placement creates geometric relationships in the tiled floor that are ratios of small integers. The alignment of the *dp* with a baseline tile corner links the vertical and horizontal discretization of the perspective tiled floor: the successive floor diagonals intersect each floor orthogonal at fractional intervals that depend on the physical size of the square tile. Therefore, changing the number of tiles at the baseline modifies the potential *dp* numbers and their placements.

[1]Controlled relaxation on the picture perspective constraints are described in § D.2.2.

Later Renaissance paintings often placed the distance point beyond the picture frame. Possible dp locations are provided outside the picture frame (Figure 5.3 (b)), at regular intervals, defined by the physical tile size. Figure 5.3 (c) shows a tiled floor using a dp inside the picture frame: the right main orthogonal is within the rightmost diamond, crossing its bottom right edge at one-third.

Effects. Perspective effects are controlled by the user, who selects the location of two points. The tiled floor has scene depth that coincides with the geometry of Talbot's planar pattern. The tiled floors in Figure 5.3 have depth geometries qualitatively different in feeling.

Figure 5.3 (a) is simplest: the frame encloses a single square pattern and the dp is on the frame. The viewer stands half the picture width from the image plane, the usual distance, looking at a floor that is proportioned, balanced and restful, a natural extension of the space in which the viewer stands.

Figure 5.3 (b) has its dp outside the frame. The intersections of the floor diagonals with the main orthogonal are lower, decreasing the foreshortening and increasing the apparent depth of the floor. A viewpoint farther than half the picture width from the image plane flattens depth, like a telephoto lens. A far away dp can bring the action claustrophobically close to the viewer.

Contrarily, in Figure 5.3 (c) a dp that is within the picture frame, close to the pvp, creates a strong contrast between deep tiles at the baseline and a crowd of indistinguishable tiles farther back. Since the viewer is close to the picture, the physical floor appears elongated. The pictorial space has a fish-eye quality, but without the hyperbolic distortion resulting from a finite-sized lens.

5.2.3 Discussion

In my framework, Talbot's planar pattern is a 2D compositional structure within which a 3D arrangement structure, the perspective floor that explicitly marks the scene depth recession, emerges without disturbing the primacy of composition. They both contribute landmarks[2] that help to blend 3D within 2D. The depth grid solves two problems.

1. The floor tiles in perspective indicate where to place objects in the scene.

2. The foreshortened lengths of the tile edges measure how to scale objects in the scene.

[2] A snapping feature to landmarks helps the user to precisely adjust the composition; see Appendix D, §D.2.

The relations between 2D and 3D blend together because:

- the principal vanishing point is at the center of a square of the planar pattern;
- the horizon coincides with a horizontal line of the planar pattern;
- the main orthogonal lines often coincide with diagonal lines of the planar pattern, and
- the inset diamonds and squares define geometric sequences with scaling factors that are ratios of small integers.

Talbot's pattern constrains the *pvp*, viewpoint rigidity being a condition for geometric harmony and relaxation of visual tension. It contrasts with the unconstrained viewpoint of the 3D virtual camera with its six degrees of freedom used in modelling software. Conversely, the grid or lattice pattern, provided universally in imaging software, emphasizes 2D composition. Its horizontal and vertical guides serve alignment and visual rhythm in the plane. But its linear rhythm interferes with the main attribute of depth: foreshortening with its geometric rhythm[3].

Talbot's planar pattern has the advantage of sharing the 2D power of a pure lattice with easy accommodation of depth geometry. For composition in addition to vertical, horizontal and diagonal guidance, its center is an advantage compared to the translational symmetry of a uniform grid. Because the picture frame breaks translational symmetry, there is a natural attraction to the center of the picture [6, p. 71], which Talbot's pattern expects and accepts, reinforcing the effect of the frame, and which is absent from the grid. Recent research correlates preferences for Arnheim's structural skeleton for the internal structure of a rectangular frame: the importance of the center, plus the diagonals and symmetrical axes in particular [83], many structural geometries that are emphasized in Talbot's pattern.

While rigid constraints apply to the tiled floor appearance in its default configuration, relaxation exists. Once this rigid configuration has been apprehended, it is fair to expand the arrangements available. Appendix D (§D.2) presents relaxed constraints on the horizon and *pvp*, which reduce the strength of the composition center: the *pvp* is no longer vertically aligned to the picture center; moving the horizon permits more room for depth arrangement.

[3] A simple, concrete example is placing the z-axis of a coordinate system at 45° when the xy-axes align with the grid. It looks like a diagonal, and the coordinate system remains planar; off 45° it fights against the grid.

5.3 Locating Points on the Tiled Floor

When constructing a scene directly on the image plane, it is sometimes necessary to know where an arbitrary world point appears on the canvas and *vice versa*. In particular, finding how an arbitrary point on one ground plane maps to the other ground plane is fundamental. Piero's diagram provides a good construction for doing so, illustrated by the drawing of a foreshortened octagonal base on the floor (Figure 4.12 (a)), using the physical floor drawn underneath the baseline of the perspective floor. Such evaluation is necessary to render arbitrary geometric volumes, regardless of the floor granularity (Appendix F) and to manipulate content within the perspective picture. (Rotating a panel requires it: Section 5.12.)

Construction. The geometric algorithm to map a point $\mathbf{Q}'$ on the physical floor to $\mathbf{Q}$ on the perspective floor requires finding the intersections of three pairs of lines (Figure 5.4). The construction is performed in two stages.

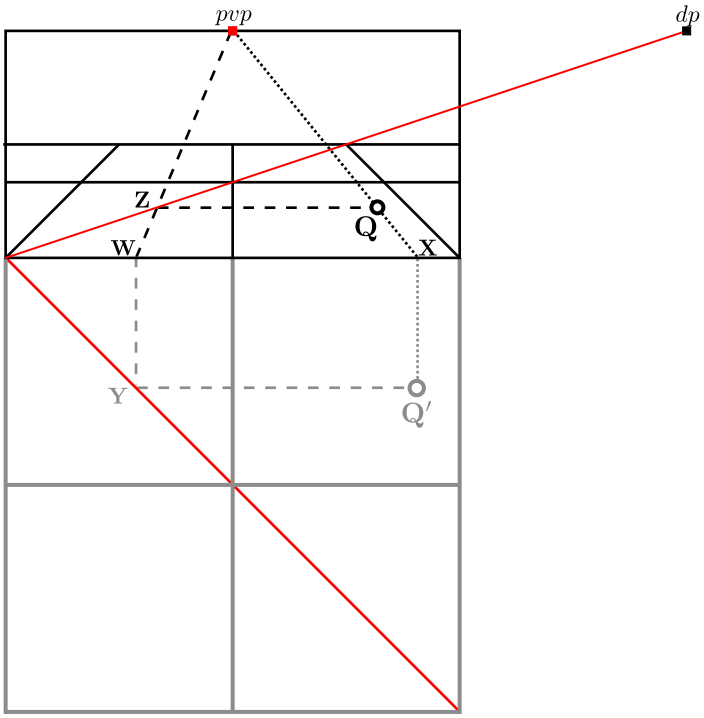


Figure 5.4: Line intersection method for transferring an arbitrary floor point between the two floors. Given a physical floor point $\mathbf{Q}'$, its perspective location $\mathbf{Q}$ is evaluated using orthogonal and horizontal lines in each floor, intersecting the baseline and Piero's diagonals.

On the physical floor three lines are drawn:

1. a vertical line through the physical point $\mathbf{Q}'$ to intersect the baseline at **X**,
2. a horizontal line through the physical point $\mathbf{Q}'$ to intersect Piero's physical diagonal at **Y**, and
3. a vertical line through **Y** to intersect the baseline at **W**.

On the perspective floor the three equivalent lines are drawn:

4. a floor orthogonal joining **X** to the *pvp*,
5. a floor orthogonal joining **W** to the *pvp*, intersecting Piero's perspective diagonal at **Z**, and
6. a floor horizontal through **Z**.

The intersection of the floor horizontal through **Z** and the floor orthogonal through **X** is **Q** the projection of the physical floor point to the perspective tiled floor on the picture.

Discussion. The construction uses lines on the physical floor and their projections onto the perspective floor, drawn as segments to keep Figure 5.4 uncluttered. The physical lines, floor horizontals and orthogonals, are edges of sub-tiles mapped into perspective in the second stage. Conceptually, the method consists of reproducing in perspective the floor lines passing through an invisible tile corner. To do so, three line orientations are used: orthogonal, horizontal, and diagonal, having their particular appearances in each space.

The horizontal line, which is the baseline, is shared by both spaces, and is thus the crucial reference for the construction. For a point directly on Piero's diagonal, floor orthogonals suffice: their intersections at the baseline defining the correspondence uniquely; the lateral position is the easiest to map because every physical floor orthogonal converges to the *pvp* in the perspective picture. How to evaluate a point's depth is less obvious. Two more intersecting lines are needed: the intersection between floor horizontals and diagonals is most practical.

In effect, Piero's diagonal plays the critical role in the construction, complementing the intersections found on the baseline. It intersects every floor orthogonal and horizontal. To determine the depth location of an arbitrary point, any line not a floor orthogonal and not a floor horizontal, can be used. Piero's diagonal is the best of these because it passes through many tile corners and is thus easy to map from one space to another, and because it spans the width and depth of the floor.

Obviously, the construction can be inverted to map from a perspective floor point to its physical equivalent. Starting with steps 6 and 7, the perspective floor orthogonal and horizontal are drawn passing through **Q**; the intersections found in reverse order determine the physical point $\mathbf{Q}'$.

This construction, which is the base for drawing in perspective, uses important principles of projective geometry: lines map to lines; parallel lines meet at the horizon or stay parallel only if parallel to the image plane; intersections remain intersections. The construction demonstrates the Renaissance break-through in geometry, as discussed in Chapter 1, § 1.2.2.

5.4 Scene Objects

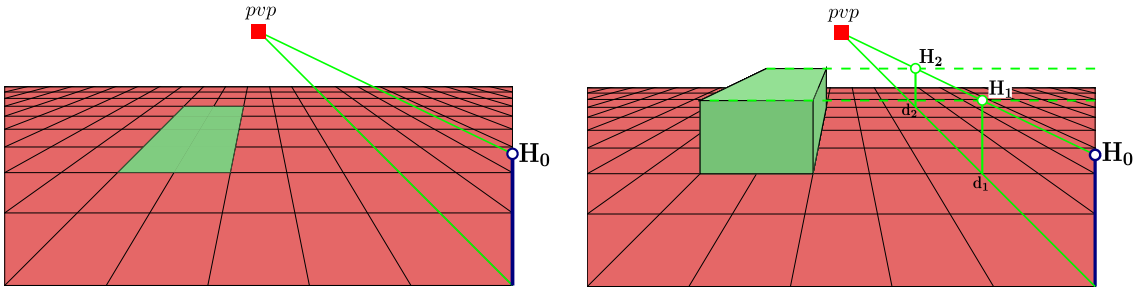
Having constructed the tiled floor defining the view of the simulated 3D scene, the scene objects comprising the picture are inserted and edited. The framework constructs their perspective appearance based on user manipulations. My framework scenes contain two types of objects. Architectural volumes reinforce the scene perspective set by the tiled floor; foreground panels, which resemble cartoons, render natural elements.

5.4.1 Architectural Volumes

Architectural volumes, including furniture, are often important elements in the scene of a picture. They form the space in which to the story of the picture occurs. They communicate the scene perspective to the viewer and define volumes from which other scene objects are excluded. Architectural elements accomplish this because of the global consistency of their visually strong edges: the lines of their forms are well-defined, long, with many parallel edges. Thus, it is usual that architectural volumes have a single perspective.

In my framework, volumes show the scene architecture schematically, simulating the absent third dimension. Editing them, the framework user explores directly on the picture the effect of scene depth: volume projections to the image plane are areas in the composition occupied by architecture. Using blocks is a simple way to define architectural elements schematically. Each block is a bounding box, presumably enclosing an architectural shape[4]. Rectangular shaped blocks have bases that are rectangular sets of floor tiles, which are elevated to bound an architectural element. Blocks can float in the air, but do so rarely.

[4]The bounding volume arbitrarily approaches the bounded object as granularity decreases and the number of volumes increases.



(a) (b)

Figure 5.5: Geometric construction for drawing a block in perspective. (a) shows the block parameters: a convex set of tiles for its base and its physical height, $\mathbf{H_0}$ (blue circle) above the baseline. (Between the green lines on the main orthogonal plane are the foreshortened heights.) (b) shows the foreshortened heights for each of the front and back depths of the block: $\mathbf{H_1}$ and $\mathbf{H_2}$ are transferred above the block base vertices at the same depth.

Thus, this section describes the construction line algorithm used for a simple block resting on the ground. All volume simulations provided by the framework are collections of blocks[5], the appearance of which respects the principal perspective of the picture.

Definition. In the physical 3D space, a block is defined by its width, its depth, its height and its location. Its projected shape on the perspective picture has location-dependent foreshortening. The lengths of the width and the depth of blocks are integer multiples of tile sizes and base vertices are at tile corners[6]. A block, representing the space occupied by an architectural volume, is specified by the user with two parameters (Figure 5.5 (a)):

1. a rectangular set of floor tiles that is its base, and
2. a height that defines its top face.

To draw a block, its six rectangular faces need to be considered:

- a base, which is a convex set of floor tiles,
- a top face, which is defined by its elevation from the base, and
- four side faces, defined by the base and top vertices.

The base is bounded by tile corners. Its side faces are easily deduced from the top face and the top face by its base and its height.

[5]Appendix D (§D.3.2) describes the hierarchies of blocks that approximate architectural structures.

[6]Because block appearance is strongly linked to floor tiles, it is most convenient to have the floor discretization such that the smallest architectural element bounding box can be represented.

Construction. Using user-specified parameters for the block the framework evaluates its foreshortened heights (Figure 5.5 (b)). Without loss of generality the following steps use the main orthogonal plane on the right of the picture.

1. Place above the bottom right corner of the frame the point $\mathbf{H_0}$, at the vertical distance of the block physical height.
2. Draw a line (an elevated orthogonal) from the *pvp* to $\mathbf{H_0}$.
3. Draw a horizontal through the front base of the block, intersecting the main orthogonal on the floor right at $\mathbf{d_1}$.
4. Draw a vertical line through $\mathbf{d_1}$, intersecting the elevated orthogonal through $\mathbf{H_0}$ at $\mathbf{H_1}$. $\mathbf{d_1H_1}$ is the foreshortened height of the front face.
5. Draw a horizontal line through $\mathbf{H_1}$, and two vertical lines from the two corners at the front of the block base. The two verticals, the orthogonal and the front of the block base define the front face of the block.
6. Repeat steps 3 to 5 at the back of the block base to define the back face of the block.
7. Draw the two top edges that recede in depth.

Discussion. The construction just described creates an axis-aligned block, but does not assume the base vertices to be tile corners (Appendix F, which presents the construction of an octagonal volume, uses such non-corner). On paper, it is executed with drawing instruments, such as straight-edge and compass, or set square and ruler. Computationally, its implementation takes advantage of the relationship between the shape and the tiled floor. Specifically, neither step 3 nor step 5 evaluate a line intersection. In step 3, for each tile corner on a same floor horizontal, i.e., sharing the same 2D y-coordinate, the tile corner on the main orthogonal is located to get the 2D x-coordinate. In step 5, the vertical distance between the main orthogonal on the floor and its elevated version encodes the foreshortened heights at all depths, the y-offset is common to all 2D x-coordinates at the same depth. Thus only the intersection between the elevated orthogonal and the vertical line is needed for each of the front and the back.

The geometric evaluation of the foreshortened heights uses one of the main orthogonal planes of the perspective floor because the larger the angle between the floor orthogonal and the vertical segment, the more accurately the intersection is located. Using a fixed region is required to maintain consistency across evaluations, regardless of the horizontal positions of the block vertices. The main floor orthogonal and its elevated version have the largest angles with the vertical lines at all canvas depths, because they lie on the most oblique orthogonal

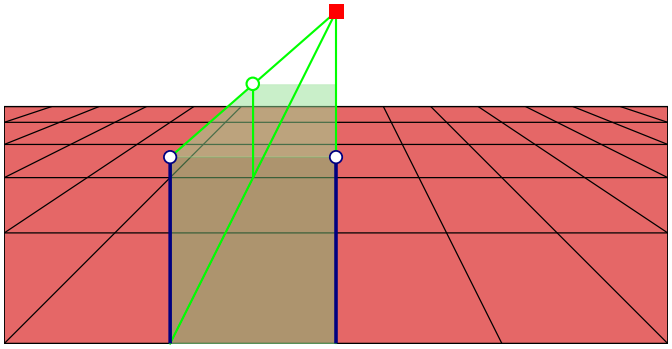


Figure 5.6: Using orthogonals passing through object edges leads to pathological cases. If one lateral face is aligned with the line of sight, for example, its elevated orthogonal and the floor orthogonal overlie one another, making it impossible to evaluate the intersection.

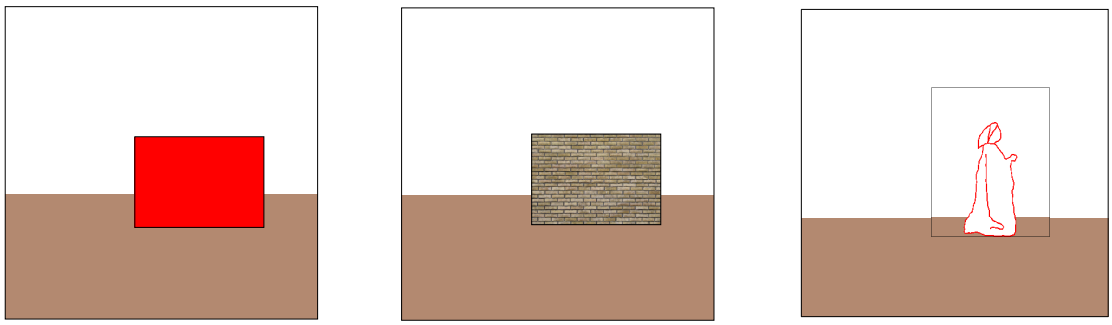
plane that lies entirely on the canvas. Thus the vertical plane through a main floor orthogonal is the ideal location to evaluate the foreshortened heights. Figure 5.6 shows an extreme example: the intersection error goes to infinity when the vertical segment and the floor orthogonal are parallel.

The block construction applies with minimal change to other solid volumes. While axis-aligned blocks coinciding with tiled floor lines simplify the algorithm—the construction and its computation are efficient due to reuse of the tiles evaluations—it generalizes as Appendix F demonstrates. However, as arbitrary volumes introduce complications for the pure 2D rendering algorithm of Chapter 7, they are left for future work (§ 9.1).

5.4.2 Panels

Architectural elements define a static spatial structure, an environment within which the other elements reside. Compositions are completed by dynamic elements caught in the midst of action. A representation of a room volume containing only furniture is dry. Life lies in the foreground figures. A composition engages by the interactions between its structural arrangement and the foreground objects: still-life, organic, animal, human and so on. Relationships among objects play out in 2D and 3D: how the two views engage is the focus of this thesis. Therefore, an effective representation for foreground objects is essential to the framework: panels are succinct 2D representations of such objects.

This section presents panel geometry: definition, location and manipulation affordances. Panels are transformed in 2D and 3D: their content riding with their geometry. Constructions that render panels as manipulations occur are given last.



(a) (b) (c)

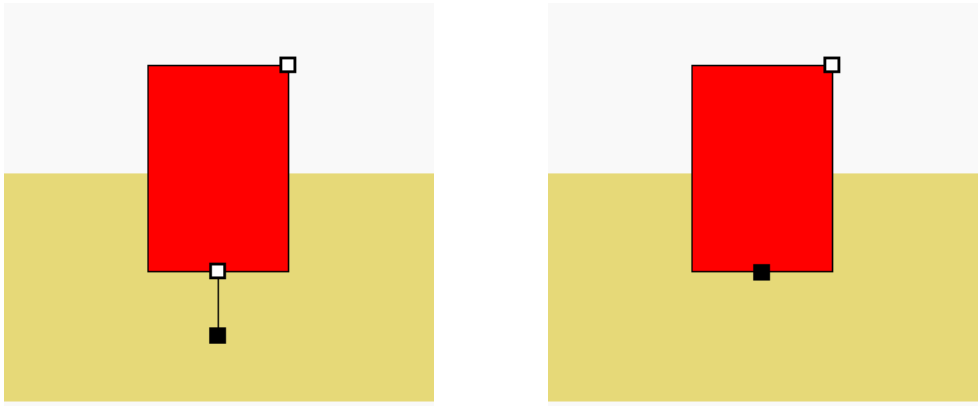
Figure 5.7: Examples of panel content. A panel is a canvas region containing an externally created content. (a) shows a panel area using the default filled colour. (b) and (c) demonstrate image content fitted in panel regions: a brick wall texture; a filled-in outline representing the silhouette in profile of a woman, where transparency is enabled.

Definition. A panel has geometric attributes defining its size and location in the scene. A panel is similar to the well-established billboard (§ 2.1.2) but its geometric attributes are formally defined. The *content of a panel* is usually a 2D projection of a 3D object, conceived and created independently of the schematic representation, in terms of a local geometry. Imitating the Renaissance cartoon, § 3.3.2, it is assumed to originate off-line as a studio study or perhaps a photograph. The framework assists the user in assembling the external content into a coherent composition, transferring the content of each onto the planar surface and into the simulated depth of the schematic representation. Because the panels are integrated into a picture with a well-defined perspective, the projection of each panel content should be generic[7] in order to maximize the recognizability of the object.

One advantage of using independent projections is that objectionable distortions of the rigid application of an unitary perspective are avoided. However, knowing the orientation of the 3D object in the final painting is important since the panel representation approximates an object best when its panel is parallel to the image plane of the final composition. The panel being a 2D approximation of a 3D object, only the 3D content facing perpendicularly the line of sight in the local depiction is visible.

In the framework a panel is a flat rectangular region, (Figure 5.7 (a)), on which is placed a 2D image. The depicted object may or may not possess volume. The panel abstraction most simply

[7]Section 3.3.2 described that the local perspective preferred for Renaissance cartoons is an independent central perspective: the viewpoint aligned to the central axis of the 3D figure being sketched; the line of sight perpendicular to the local canvas. Thus, local foreshortening is present and not distorted and will adhere to the central foreshortening of the principal perspective of the picture the external content is inserted into.



(a) (b)

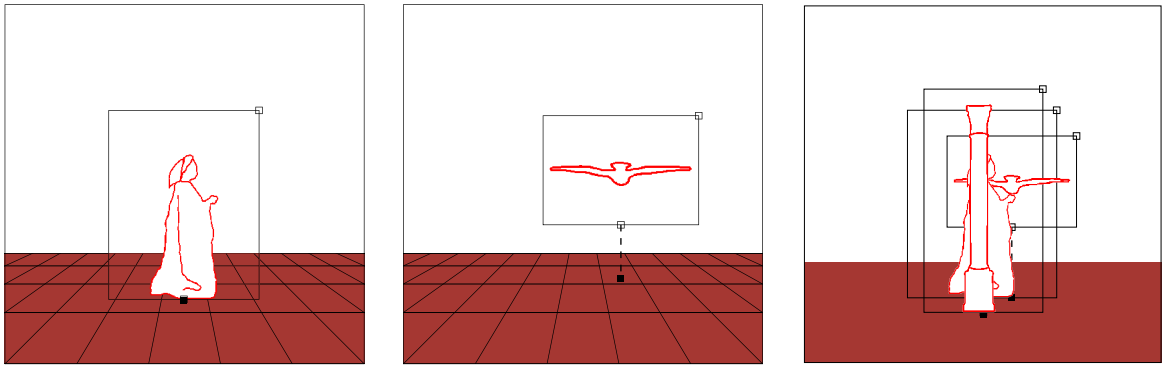
Figure 5.8: A panel has an origin and an attachment point that specify the 3D location of the object represented. The attachment point disambiguates the panel placement with respect to the ground plane: its position on the panel central axis defines a depth. Using it, users stipulate if an object is in contact with the ground or if it is floating above it.

depicts flat content as a visual texture (Figure 5.7 (b)). When content has volume, a filled-in outline of the locally projected silhouette is sufficient as long as the panel is small compared to the picture, which is often adequate for human shapes or organic ones (Figure 5.7 (c)). Partially transparent panels show precise occlusion, as demonstrated in Figure 5.9 (c).

Position. In my framework application, a set of images beside the canvas (Figure 5.1) provides ready-made panel content. Images are transferred to panels using drag-and-drop. The user then adjusts the two parameters that fully encode the object placement within the scene: its origin and its attachment point (Figure 5.8).

The panel's *origin*, at the center of its base specifies the location of the panel in the 2D space of the image. However, the panel's origin does not specify when its place in the drawing sequence, which determines occlusions that occur in 3D space. The panel's *attachment point*, located on the vertical mid-axis of the panel, defines the object's position on the tiled floor. The vertical offset between the attachment point and the origin determines the height of the panel above the ground plane. Adjusting the attachment point changes the depth of the panel and thus when it is drawn relative to the other objects of the scene.

Figure 5.8 illustrates two relationships between the attachment point and the origin, corresponding to two distinct interpretations of a single panel: in (a) it hovers over the ground; in (b) it rests on the ground, the attachment point coinciding with the origin.



(a) (b) (c)

Figure 5.9: (a) a woman at the composition center is represented by her silhouette on a panel: its attachment point connects to a central floor tile on which she stands. (b) a bird flying above the ground, its panel attachment is below the panel origin. (c) those panels are combined with one representing a column: the attachment points define the occlusion among them.

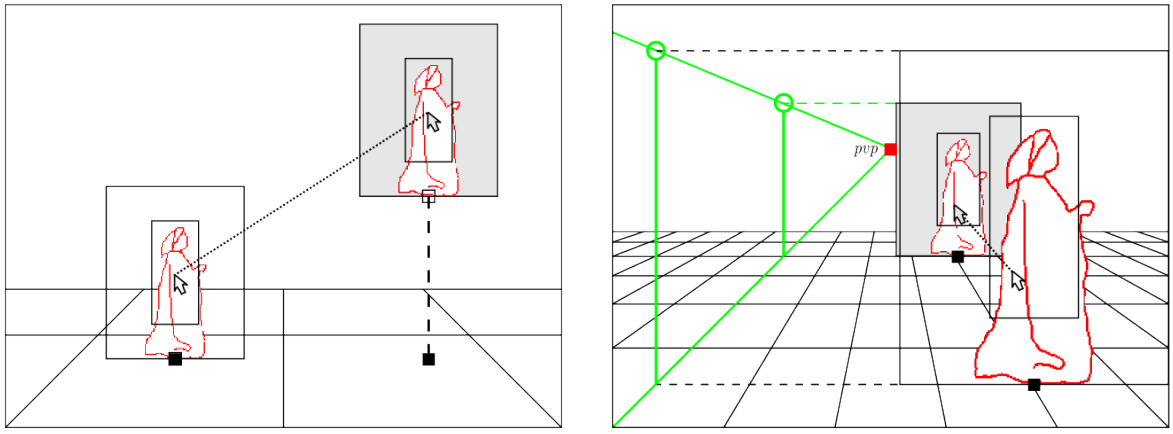
The attachment point places the panel at a specific depth on the perspective tiled floor, the depth of the tile that owns it. Figure 5.9 shows attachment points with the tiled floor visible. The attachment points determine the depth ordering that correctly renders the scene: in the column occludes the woman, and both occlude the bird.

Panel Affordances. The user moves, sizes and rotates panels to compose and arrange the foreground of a schematic representation. Interactive feedback from the framework updates the 2D geometric appearance of the scene using construction line evaluations.

A panel is moved either in 2D, translated in the frontal plane to alter the 2D composition (Figure 5.10(a)) or in the scene depth to change the 3D arrangement (Figure 5.10 (b)). Construction line evaluations are needed to scale the panel appearance as its depth changes. Independently, the physical size of the panel is scaled uniformly about its center, Figure 5.11 (a), the origin and attachment point move; about its origin, Figure 5.11 (b), staying at the same location with respect to the floor.

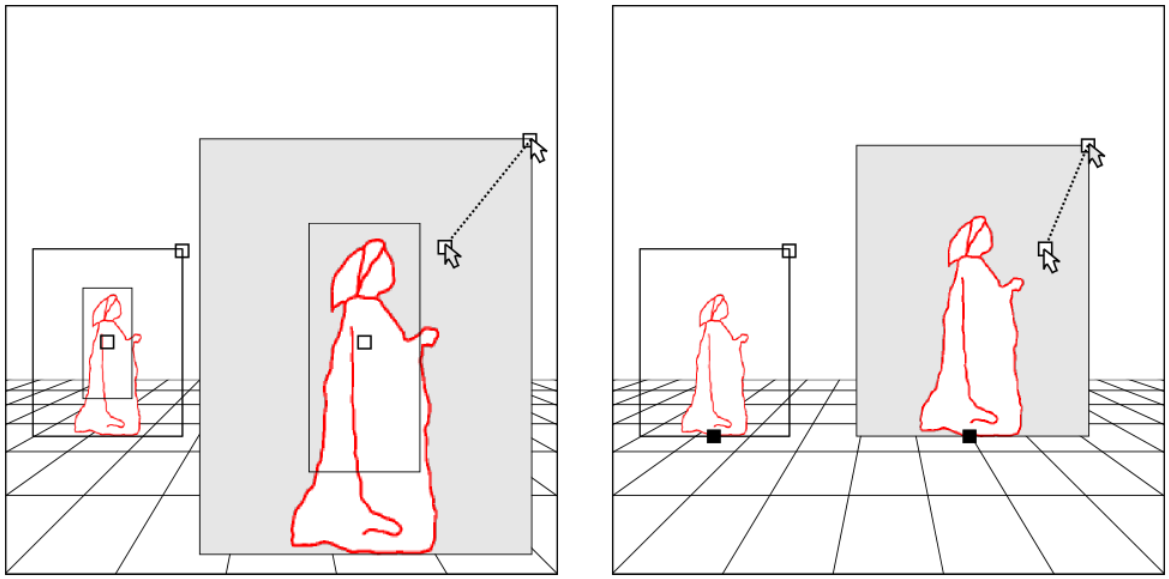
The user can also rotate a panel around its vertical mid-axis to tune its orientation. Two view directions define the rotation: one is orthogonal to the image plane, the other is orthogonal to the panel, and is usually the view direction with respect to which the panel content was prepared. By default the two are parallel; they diverge as the panel is rotated.

Figure 5.12 compares the effects of rotating by a pair of angles at different lateral positions.



(a) (b)

Figure 5.10: (a) 2D motion in the frontal plane, the location of the panel origin is changed, the rest following along. (b) 3D-like motion in the scene depth, the panel is pushed closer to *pvp* (along the line parallel to the line of sight): the panel height foreshortens appropriately.



(a) (b)

Figure 5.11: Two types of scaling, 2D and 3D, change the size of panels. In (a) the center point (the unfilled middle square) stays fixed; whereas in (b) the panel origin stays at the same location. On each image the panel on the left is the reference showing the initial size before the manipulation. (It was also moved laterally in this figure for clarity.)

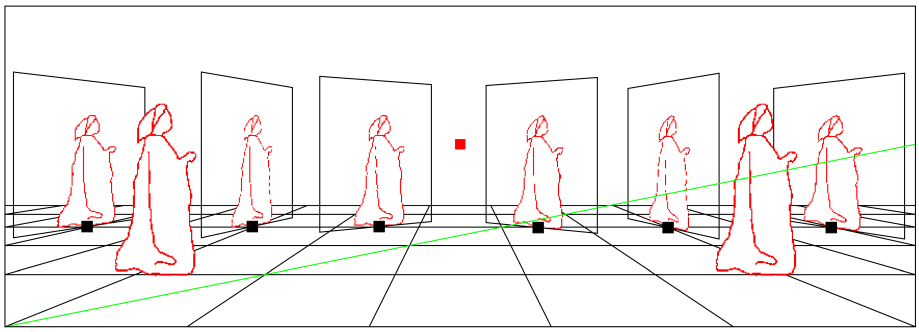


Figure 5.12: On the back row, each panel is rotated by an angle: from left to right 70°, 70°, 30°, −30°, −70° , −70°. A consistency check of the construction applies: for a same rotation magnitude the base and top edges of the panel boundary converge on the horizon at the same distance from the *pvp*: $vp_{|\theta|}$. On the front row the panels are in the default orientation.

Figure 5.12 illustrates that extreme rotations are ill-advised as they introduce wide-angle distortions, which increase with the lateral distance of the axis of rotation from the *pvp*. The two lateral positions, where a rotation of |70°| is used, show the evolution of the distortion: the farther away from the *pvp* the panel is, the wider the panel becomes. Indeed, the rotation distortion increases with angle size and with lateral distance from the *pvp*. Several constructions are required to change the 2D appearance of a rotated panel in accord with the principal perspective of the picture: the local depiction preserving geometric consistency with the picture depth. If distortions were to be minimized, an alternate construction exists: rotating the panel vertical axis aligned to the *pvp* and repositioning the result to the lateral position[8].

Constructions. The geometric constructions that implement the above panel affordances are now presented. Some changes in appearance are straightforward, not requiring construction line evaluation. Others depend on the principal perspective of the picture. One such construction resizes a panel translated in depth (Figure 5.10 (b)) and is already known: it is the one used to elevate a block (Figure 5.5). Three more constructions are required. A pair of them determines the physical measures that have been modified on-the-fly by user initiated manipulations: the constructions to retrieve the panel's height and width follow on Piero's diagrams, Figure 4.12. A third construction evaluates the foreshortening of the boundary of a panel that is rotated around its vertical mid-axis. Having added the perspective affordance, I devise its construction from concepts learned from other geometric constructions.

[8]This consistent perspective construction follows from the one described below, and has the advantage of adhering to the global scene geometry.

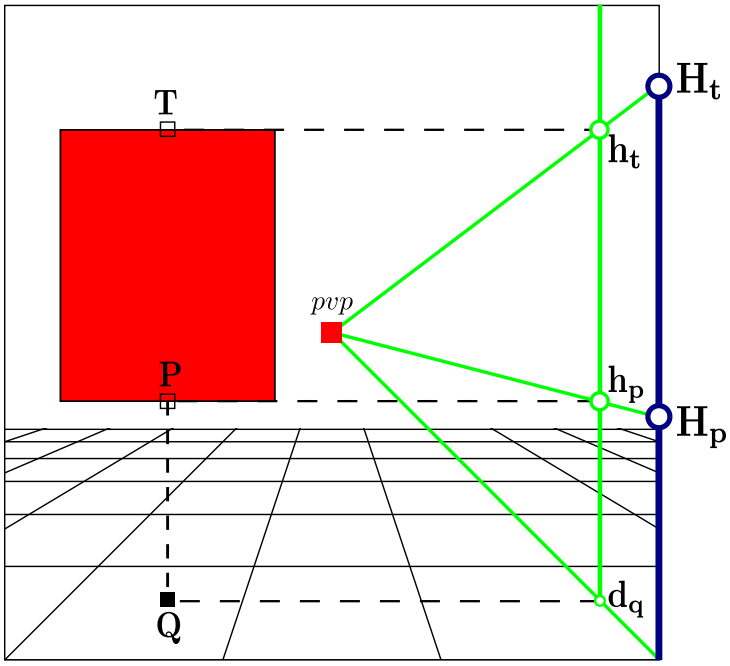


Figure 5.13: Construction lines to evaluate physical heights from their foreshortened appearance in the perspective picture. The evaluation is done on the main orthogonal plane with lines parallel to the floor horizontals using the attachment point locating the heights in depth.

Figure 5.13 shows construction lines that evaluate the physical heights after a user's adjustments: both the panel origin, **P**, and its top boundary edge, on which the point **T** lies, have been moved so that their physical heights above the panel attachment point, **Q**, are unknown. The right main orthogonal plane is again used for the construction. The steps below describe in words the construction to determine the physical height of the origin only, **P**. (The construction to derive the physical height of **T** is similar: steps 3 to 5 are modified replacing **P** and the subscript **p** by **T** and **t** respectively).

1. Draw a horizontal line through **Q** to intersect the main orthogonal at $\mathbf{d_q}$.
2. Draw a vertical line through $\mathbf{d_q}$.
3. Draw a horizontal line through **P** to intersect the vertical line through $\mathbf{d_q}$ at $\mathbf{h_p}$.
4. Draw an elevated orthogonal from *pvp* through $\mathbf{h_p}$ to intersect the frame edge at $\mathbf{H_p}$.
5. The vertical distance from the baseline to $\mathbf{H_p}$ is the physical height separating the panel origin from the floor.

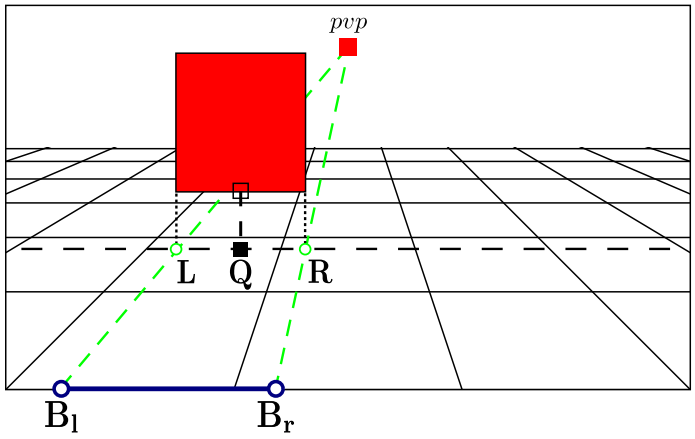


Figure 5.14: Construction lines to evaluate a physical width from a foreshortened length. The evaluation is done at the baseline using lines parallel to the floor orthogonals passing through the attachment points of the horizontal segment end-points.

Since panel manipulations always affect the panel width and height sizes uniformly, the width is easily derived from the physical height. However, as an introductory stage to compute the appearance of a rotated panel and for the sake of completeness calculating physical distances from foreshortened widths is valuable. Figure 5.14 shows a panel floating above the floor, with attachment point **Q**. The following steps derive its physical width.

1. Draw a horizontal line through **Q**.
2. Draw the vertical projecting the left edge of the panel boundary to intersect the horizontal through **Q** at **L**.
3. Draw a line from the *pvp* passing through **L** to intersect the baseline at $\mathbf{B_l}$.
4. Redo steps 2 and 3 for the right edge of the panel boundary. **R** is the intersection with the horizontal through **Q**; $\mathbf{B_r}$ is the intersection with the baseline.
5. The length $\mathbf{B_l B_r}$ is the physical width of the panel.

To evaluate the physical width of a panel, its boundary projection onto the floor is essential. For an unrotated panel the boundary horizontal endpoints on the floor transferred to the baseline give its physical width. Were the projected segment onto the floor not-horizontal however, this construction is insufficient, as demonstrated next.

A construction for rotation is now described. For a panel with attachment point **Q** and origin **P**, the construction, which rotates the panel by θ around its vertical mid-axis **QP**, is broken into three stages.

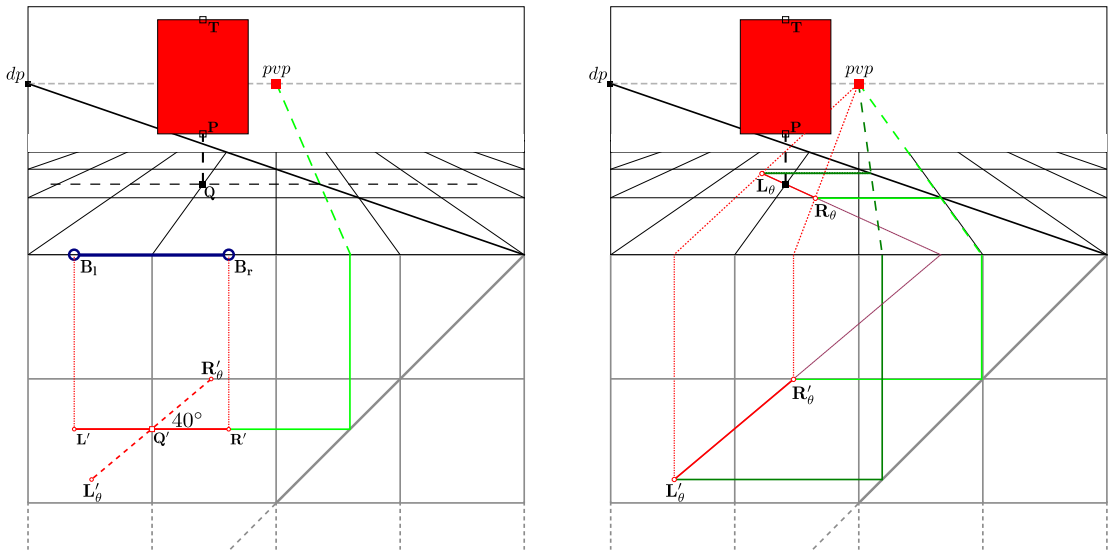


Figure 5.15: Construction lines to evaluate the foreshortened panel width on the perspective floor for a panel vertically rotated by $\theta = -40°$. (a) the unrotated panel width is mapped to the physical floor and rotated by $-\theta$; (b) the physical rotated segment is mapped back to its location on the perspective floor.

Figures 5.15 (a) and (b) illustrate the evaluation of the foreshortened panel width on the perspective floor, the panel boundary is yet unrotated; Figure 5.16 shows the construction of the rotated panel boundary from the foreshortened segment onto the perspective floor. The construction steps for each figure are as follows.

1. Figure 5.15 (a): the unrotated width is moved to the physical floor to be rotated by $-\theta$.
   (a) Map $\mathbf{Q}$ to $\mathbf{Q}'$ using the inverse construction of § 5.3[9].
   (b) Determine at the baseline the panel physical width of the panel, $\mathbf{B_l B_r}$ (Figure 5.14).
   (c) Draw vertical lines from $\mathbf{B_l}$ and $\mathbf{B_r}$ on the physical floor.
   (d) Extend the horizontal line through $\mathbf{Q}'$ to intersect the vertical lines at $\mathbf{L}'$ and $\mathbf{R}'$ respectively.
   (e) Rotate $\mathbf{L}'\mathbf{R}'$ around $\mathbf{Q}'$ by $-\theta$ to $\mathbf{L}'_\theta\mathbf{R}'_\theta$.
2. Figure 5.15 (b): the physical rotated segment is mapped to the perspective floor.
   - Map $\mathbf{L}'_\theta$ and $\mathbf{R}'_\theta$ to $\mathbf{L}_\theta$ and $\mathbf{R}_\theta$ respectively, using the construction of § 5.3.

[9]On each floor, an horizontal and a pair of lines parallel to the floor orthogonals are used with Piero's diagonal and the baseline to determine $\mathbf{Q}'$. For Figure 5.15 (a), only one orthogonal is needed as $\mathbf{Q}$ is on a floor orthogonal.

3. Figure 5.16: the foreshortened rotated panel boundary is determined.
   (a) Draw a line $\mathbf{L}_\theta\mathbf{R}_\theta$ to intersect the horizon at vp_θ.
   (b) Draw a line from vp_θ through the panel origin $\mathbf{P}$.
   (c) Draw a line from vp_θ through the center of the panel top edge $\mathbf{T}$.
   (d) Draw the two vertical lines through each of $\mathbf{L}_\theta$ and $\mathbf{R}_\theta$.
   (e) Their intersections with $vp_\theta\mathbf{P}$ and $vp_\theta\mathbf{T}$ define the rotated panel boundary.

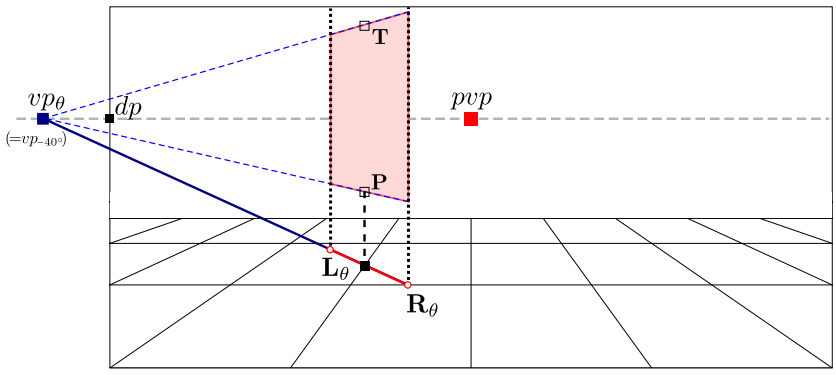


Figure 5.16: Construction of the boundary of a vertically rotated panel from its base $\mathbf{L}_\theta\mathbf{R}_\theta$ on the perspective floor. $\mathbf{L}_\theta\mathbf{R}_\theta$ intersects the horizon at vp_θ. The bottom and top edges of the panel through the two invariant points, $\mathbf{P}$ and $\mathbf{T}$, also converge to vp_θ. The boundary of the rotated panel is then determined by their intersections with the verticals through $\mathbf{L}_\theta$ and $\mathbf{R}_\theta$.

The construction described is not unique; variations exist. For the third stage, one could determine the corners of the rotated panel by using the construction that elevates floor points (§ 5.4.1). Instead the chosen construction uses the convergence property of points on the horizon for parallel lines. The two parallel edges of the panel and their projection onto the floor, $\mathbf{L}_\theta\mathbf{R}_\theta$, in the perspective image converge to a single point on the horizon, vp_θ. While the vp_θ construction has precision that depends on θ and panel lateral position, and is impractical for small rotations, it suggests more possibilities for simulating 3D with construction lines. The construction may not be the best choice, but it illustrates an efficient check on the obvious variant: for artists constructing on the canvas, it checks for drawing errors; for programmers it checks for implementation bugs.

Figure 5.15 (b) illustrates a second consistency check, resulting from Piero's juxtaposition of the spaces at the baseline, the single line on the ground that is common to the two spaces. Therefore, the rotated line on the physical floor, $\mathbf{L}'_\theta\mathbf{R}'_\theta$, must meet the baseline at the same point as its perspective equivalent, $\mathbf{L}_\theta\mathbf{R}_\theta$.

It is essential to use a rotation by $-\theta$ on the physical floor to evaluate a panel rotation θ in the perspective image: negating the angle undoes reflection in the baseline when mapping the physical floor into the perspective one, shown by the Piero's inverted relation[10]. In term of correctness, it is worth mentioning that rotation of $\pm 90^{\circ}$ can be constructed using the method described as long as **Q** is not vertically aligned with the pvp. When the panel's attachment point is vertically aligned to the pvp, elevating the points for the third stage is, however, robust. However, extreme changes of orientation would normally be avoided as they do not make sense for planar representations of objects.

Given the rotated panel boundary, it is trivial to implement the corresponding transformed image of the panel content as described in Section D.3.1. On a computer the algebra is straightforward because the transformation is projective and corresponding conceptually and mathematically to texture mapping: lacking such automate processes Renaissance artists transformed a grid drawn on a cartoon and used it as a guide for hand drawing such 3D-like transformations, similar to their practice in drawing anamorphic distortion [65, p. 210].

Discussion. This section described panels and methods available to framework users for simulating 2D and 3D-like manipulations of flat representation of local objects. The framework couples 2D and 3D interactions that modify the appearance of panels on the canvas, taking advantage of flatness and simulated scene depth.

The constructions that modify panel appearance use three landmarks: the main orthogonal plane, the baseline and the horizon. The descriptions of the constructions include techniques that maintain geometrical correctness, specifically checking convergence at the baseline and at the horizon. These possibilities result directly from properties of projective geometry and its relationship to Euclidean geometry, which are instructively combined in Piero's juxtaposition of floor spaces.

It is relevant to note that the dp does not directly affect most manipulations: the constructions that affect position, scale, physical height and width for the unrotated panel appearance, do not use the dp. Indirectly those transformations and physical evaluations have a relation to the dp as everything has a meaning in term of the fraction of the constant tile physical size, whose perspective appearance is defined according to the dp distance to the pvp.

However, rotating a panel is different. The foreshortened appearance depends directly on the position of the dp because the distance of the viewer alters the field of view. More or less

[10] Another visualization to comprehend the sign inversion of the angle is to notice that Piero's physical floor on the canvas is, in effect, seen from below.

noticeable perspective distortions are introduced depending on the panel's central alignment compared to the pvp: a farther dp increases stability in the appearance of a rotated panel. In the perspective floor, Piero's diagonal that crosses the canvas from left to right, meets the horizon at the $dp = vp_{45^\circ}$: between it and the pvp are vp_θ for $\pi/4 \leq \theta \leq \pi/2$; beyond the dp, to ∞ on the horizon, are vp_θ for $\pi/4 \geq \theta \geq 0$. Changing the dp location contracts or expands the segment where the parallel lines of different orientations converge on the horizon.

5.5 Shadows

This section describes artists' constructions that produce projected shadows in my framework: algorithms evaluate the double projection geometry on the canvas. Depicting projected shadows in 2D is similar to calculating and intersecting shadow volumes in 3D. Geometric constructions evaluate in between pairs of orientation of shadowing and shadowed planes. The construction lines to recreate the 3D reality using 2D projective geometry demand attention: problematic cases exist, owing to construction lines passing through the illuminant's center of projection.

5.5.1 Scope

Shadows are included in the framework, concentrating on the effects artists use most commonly in practice. Artists normally light the scene with an illumination in roughly the location of the artists so that faces are not in shadow. They usually prefer a source of light that is large, behind them and enough off to the side for differential lighting to bring out the three-dimensionality of the face[11]. Projected shadows affect visual perception only generally providing the illumination context within which human vision gathers information in the 3D world. In depiction, artists use an efficient compromise: the shadow's overall appearance needs to be convincing without the details needing to be physically correct [13].

Like artists, my framework creates projected shadows using 2D objects. For volumes, especially block elements, the projected shadows rendered are the real ones, constituted from a few flat faces with exact geometry. However, for panels, which fill the foreground, projected

[11] A few artists like to put the sun in the picture: Turner and Le Lorrain come to mind. Turner handles this potentially problematic illumination with Turneresque haze, wind and clouds so that the light seems to come from everywhere. Le Lorrain places his sun low in the sky giving strongly coloured yellow or bronze illumination. His scenes are wide with open centre, thus the viewer sees illuminated fronts on both sides of the pictures.

shadows differ in detail from silhouettes of 3D objects. The projected shadows generated by the framework only approximate those produced by full 3D object geometries.

In addition, only some projected shadows are included. Generalizing the artist's rule of thumb process, in which local adjustments are easy to make, is cumbersome. Cole's manual describes many particular cases, different surface configurations under specific illumination: shadows on the ground, on vertical planes, on stairs cast by simple objects standing on the ground or dangling from a ceiling [17]. In the framework only shadows falling on critical surfaces are included, those falling on the ground, and those projected by panels onto block faces. Other shadows, such as shadows of panels falling on each other, are omitted as they intermix with shading[12], which is not addressed in the framework.

5.5.2 Definitions

On the canvas, the geometry of projected shadows has combinations of two surfaces and two types of illumination rays, which are defined as follows.

- **The light source** is the location of the illumination. It indicates where the 3D light source projects onto the canvas, specified by vertical and horizontal placement.
- **The light vanishing point**, lvp, is the light attachment point, indicating the light's projection onto the ground plane.
- **The shadowing object** is the surface that blocks the light.
- **The shadowed plane** is the surface on which the shadow is cast.
- **Light rays** are lines from the light source to the shadowed plane in the neighbourhood of the shadowing object.
- **Ground rays** are lines on the shadowed plane from the lvp, passing through the intersection of the shadowing object and the ground.

The position of the light source on the canvas is insufficient to define the illumination position in the 3D scene. The light vanishing point defines its position with respect to the ground, resolving the ambiguity. The user specifies both the light source on the canvas and its vanishing point, lvp, on the vertical axis running through the source light, which indicates its depth on the ground plane. The lvp is called the light vanishing point by analogy to the scene principal vanishing point: it is on the ground plane and is a point of convergence for shadow geometry lines.

[12] Artists include them as part of the colouring of surface: Willat's *denotation system*.

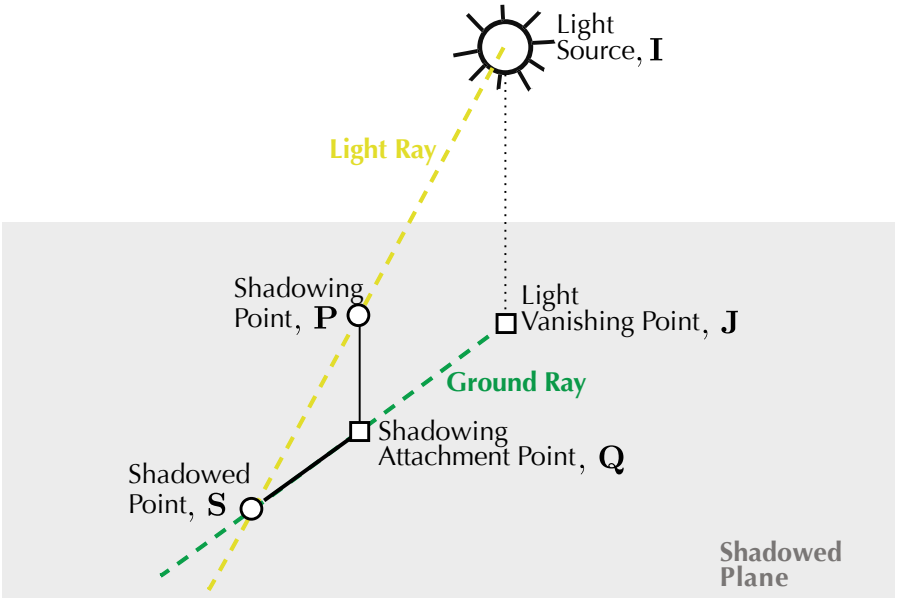


Figure 5.17: The ground shadow geometry of a general point. The light ray, from the light source point to the shadowing point above the ground, intersects the ground ray, from the *lvp* to the shadowing attachment point: this intersection is the location of the shadowed point.

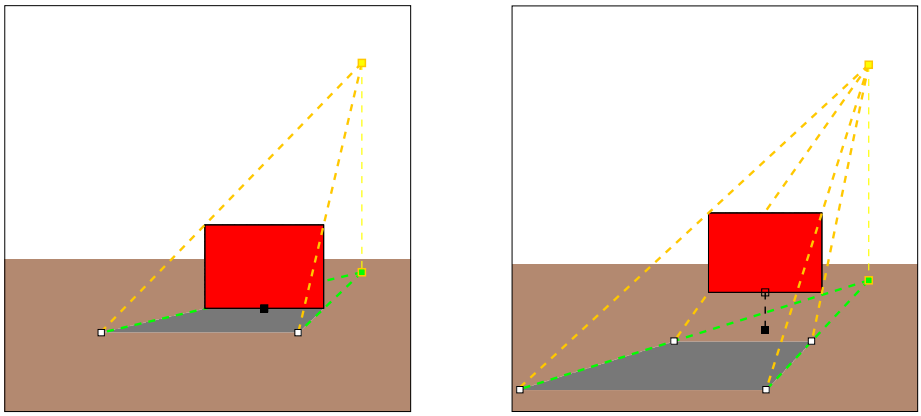
The illumination representing a low sun has its vanishing point on the horizon, at infinity on the ground, for the shadow effects at sunset or sunrise, both points (the *lvp* and the source) being within the picture frame or close by; when the sun is high or lighting comes from the side of the scene on the picture, its representation on the canvas needs to change to a direction.

5.5.3 Ground Shadows

The projected shadows that fall on the ground are the most important shadows: each anchors an object's position in the scene depth. Their geometric construction is simple. The shadowed plane is the ground.

Single Point. Figure 5.17 shows that 2D geometric calculations are sufficient to determine for a general point **P** with attachment point **Q**, the location of its projection, **S**, onto the ground plane with a light source at **I** *lvp* at **J**. The construction steps follow.

1. Draw the light ray from **I** passing through the shadow casting point, **P**.
2. Draw the ground ray from **J** passing through the attachment point, **Q**.
3. **S**, the intersection of **IP** and **JQ**, is the location where the point's shadow falls on the ground plane



(a) (b)

Figure 5.18: In (a) the attachment point on the base of the surface creates a ground shadow that is connected to the shadowing wall. In (b) the attachment point below the surface base creates a cast shadow that has no point of contact with its shadowing wall.

However, artists rarely realize the shadow of a single point for its own sake, but treat together a collection of points inside the silhouette of an opaque object. Applying the single point construction enough times it is a guide for approximating a cast shadow (Figure 4.13). Cole's examples include projected shadows cast by an arbitrarily oriented wall, which requires the single point construction to be applied at least twice. Further, Cole's example of the shadow produced by an arch requires many more points: my framework imitates such a result using the panel representation. Each is described below.

Opaque Polygon. Figure 5.18 presents two simple geometric constructions creating the shadow cast by a vertical wall positioned parallel to the image plane. In (a) the base of the wall is on the ground plane and its corners coincide with their own shadow points, the evaluation of the two shadow points of the wall top corners is sufficient. In (b) defying gravity the wall hovers in mid-air and the evaluation of four shadow points is needed[13].

Figure 5.18 illustrates the redundancy that commonly occurs in the ground shadow projection of a planar polygon: its orientation with respect to the image plane and its location with respect to the ground plane matter, reducing the complexity of the general construction. In fact, the ground shadow of a general n-sided polygon arbitrarily placed requires the point construction n times, totalling a maximum of $2n$ rays used[14].

[13]Four light rays through the wall corners are needed, since they are distinct from their attachment points; yet only two ground rays are used.

[14]m light rays passing through each vertex point that does not coincide with the vertex attachment point and m' ground rays passing through each distinct attachment point of the n vertices. (Both m and $m' \leq n$).

The critical aspect of the ground shadow construction is that artists intuitively know the position of the shadowing object relative to the ground. Knowing this they intuitively place the straight edge correctly. For the framework the object attachment point determines the location of the cast shadow. User adjustment of the object attachment point controls the shadow appearance, the framework derives the attachment points of the polygon corners according to the polygon's orientation. The interaction of two surfaces, the shadowing surface and the shadowed one, in the presence of illumination produces a shadow. The two surfaces determine the shape of the shadow. The light rays intersect the shadowing and shadowed planes in an order that depends on the position of the illumination. Sometimes rays must be extended beyond where the shadowing object intersects the ground.

Such a point-by-point construction is not efficient for creating the ground shadow of a flat object of arbitrary shape, which may contain detailed curves. The panel representation, however, enables a convenient simulation based on the mapping of its quadrilateral boundary. The mechanism that simulates the projected shadows cast by complex silhouettes, imitating the shadows painted by artists, is presented next.

Panel Collineation. The framework includes shadows produced by partially transparent panels, rectangular regions containing complex 2D content, the equivalent of artists' cartoons. Shadows of opaque faces are not particularly interesting and take little advantage of a computer. Contrarily, panel content casts significant shadows that reinforce the shape of the objects casting them. While precise shadows are impossible because only one silhouette is available, my framework's shadows are comparable to those created by artists.

To produce the ground shadow cast by a partially transparent panel, the framework relies on the line construction geometry used for the opaque wall (Figure 5.18): Figure 5.19 demonstrates the line construction with a cast shadow of compelling complexity. This simulation makes more use of the power of a computer. From a few line intersection evaluations, the inside transfer is automated, minimizing the constructive computation and simplifying Cole's methods, which apply the point-by-point construction many times, as illustrated in Figure 4.13 (c).

Lights are centers of projection for projected shadows and like the tiled floor they have vanishing points. Light vanishing points are also centers of projection. Therefore, given the ground and light rays four intersection points on the ground of the quadrilateral boundary of the shadowing panel face, the quadrilateral shape they form on the shadowed ground defines a perspective transformation. Thus, from the mapping between the panel boundary corners and their equivalent projected on the ground plane, a 2D projective transformation, or collineation,

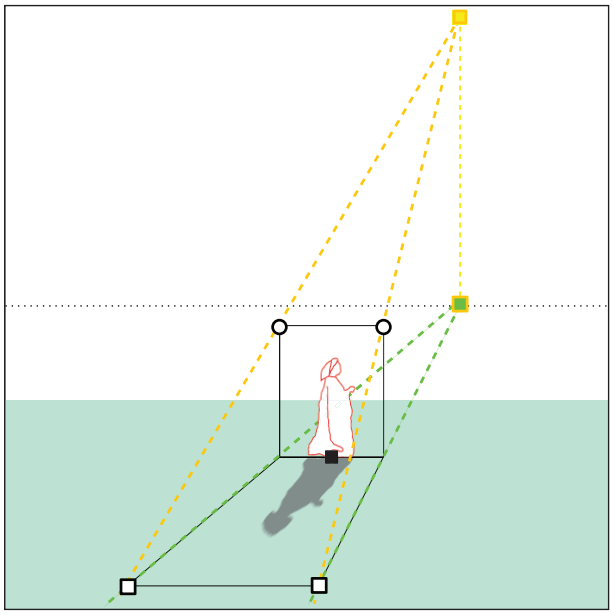


Figure 5.19: The ground shadow created by a panel illuminated by a sun. The construction for evaluating the projection of the panel boundary is simple (Figure 5.18). A collineation then transforms the panel content mask to simulate the ground projected shadow.

is derived, as for the appearance of content on a rotated panel (Section D.3.1). To produce the projected shadow shape, a mask image is created from the content of the partially transparent panel. The collineation then warps the mask into the projected quadrilateral. The image of the transformed mask is painted on the ground: 2D texture mapping automatically reproduces the artist's transfer of a complex shape into its shadow using a grid (see Section D.4).

In some cases the framework's projected shadow might differ slightly from the result produced by the equivalent 3D scene, where the 3D object content is available. However, artists have found such shadows acceptable for centuries, as they conserve all the qualities that are important for vision. As a canonical view is often used for the orientation of the 3D object depicted in the panel content, the main features identifying the object are in the local projection and in its shadow. The shadow excludes extraneous object details that would disturb the viewer appreciation of the composition.

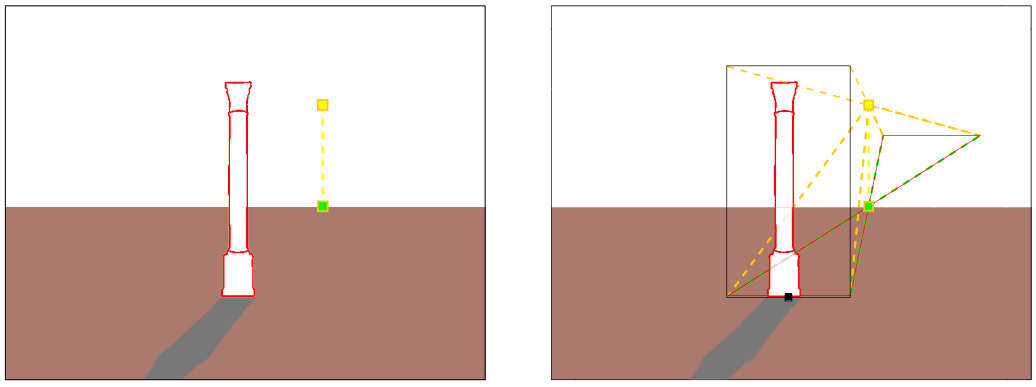
Special cases. The above geometric construction is simple and intuitive but it is not robust in the general case. For example, when the boundary of a panel has one vertical edge aligned with the light axis, the intersections between the light and ground rays for that side of the panel,

are undefined (similar to the block elevation shown in Figure 5.6). An artist immediately notices such problems and uses equivalent oblique rays or picks the most oblique pairs of rays to find intersections that are extrapolated horizontally. The framework version of the artist's construction always uses the most oblique pairs of rays for unrotated panels so as to provide robust results.

Most significantly, problems occur when the shadow of a panel extends beyond the picture frame (Figure 5.20 (a)). Sometimes clipping the ground shadows to the canvas frame is sufficient. However, this remedy does not always work. Ground shadows can be infinitely long, extending through the light vanishing point (Figure 5.20 (b)). Because the light is low relative to the panel height, light rays through the panel's top vertices intersect the ground rays on the other side of the light—behind it.

To handle situations in which infinitely long shadows occurs my framework reproduces the practice of artists who are intuitively guided by the canvas appearance. The shadowing boundary is clipped to the region that produces the visible shadow within the picture. The transparent yellow region in Figure 5.21 is the visible shadow of the column to the image plane. A similar situation occurs with the horizon when a low illumination is closer to the baseline compared to the shadowing object.

While ground shadows using only object attachment points and lvp are independent of the picture perspective, shadows on other surfaces often interact with it. For projected shadows on most other shadowed planes, the perspective captured in the tiled floor is needed.



(a) (b)

Figure 5.20: When the illumination is low the construction of Figure 5.19 might fail owing to the height of the shadowing height relative to a low illumination source. Artists correctly depict the shadow with ease, considering only its appearance to the baseline.

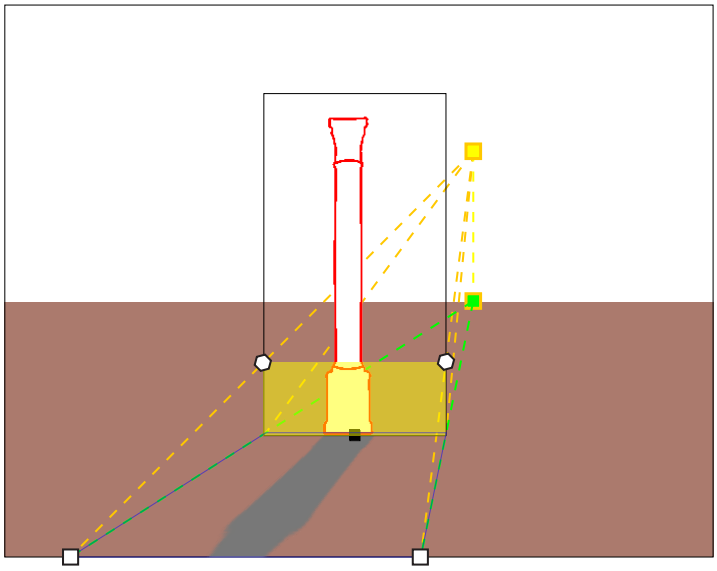


Figure 5.21: Geometrical construction for a ground shadow when clipping is required: the baseline intersection points of the ground rays for the panel vertical edges are used to draw only the part of the shadow behind the image plane.

5.5.4 Other Shadows

Like shadows falling on the ground, shadows falling on other plane surfaces are delimited by straight edge constructions. The framework renders the shadows using the masks of panel content that are projectively transformed. Displaying the construction lines that link the shadowing element and shadowed surface in relation to the illumination source position helps the framework's user to visualize the lighting geometry and thus to control the shadow appearance. This section describes shadows projected onto projected planes, which are generated as part of the scene composition.

On Elevated Ground Planes. When the shadowed plane is not the ground, but a plane parallel to it, most often elevated above it, the construction is an adaptation of the earlier construction (Figure 5.19). The ground rays are drawn on the elevated plane. Thus the objects' attachment points and the light vanishing point on the elevated plane must be determined. Figure 5.22 shows such a shadow projected on an elevated ground, as the panel casts a shadow on the top face of a block: on the light axis the elevated light vanishing point (the circle) is used for the construction.

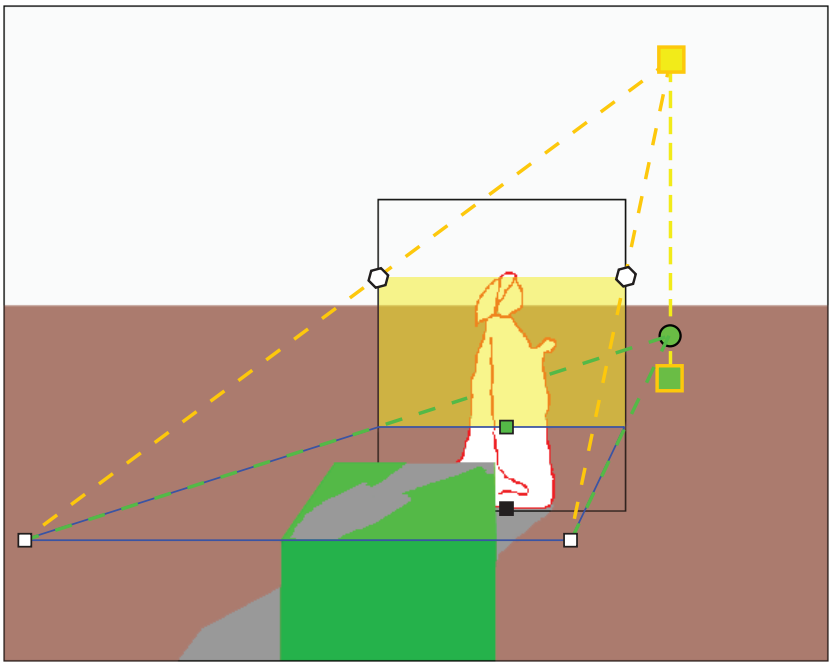
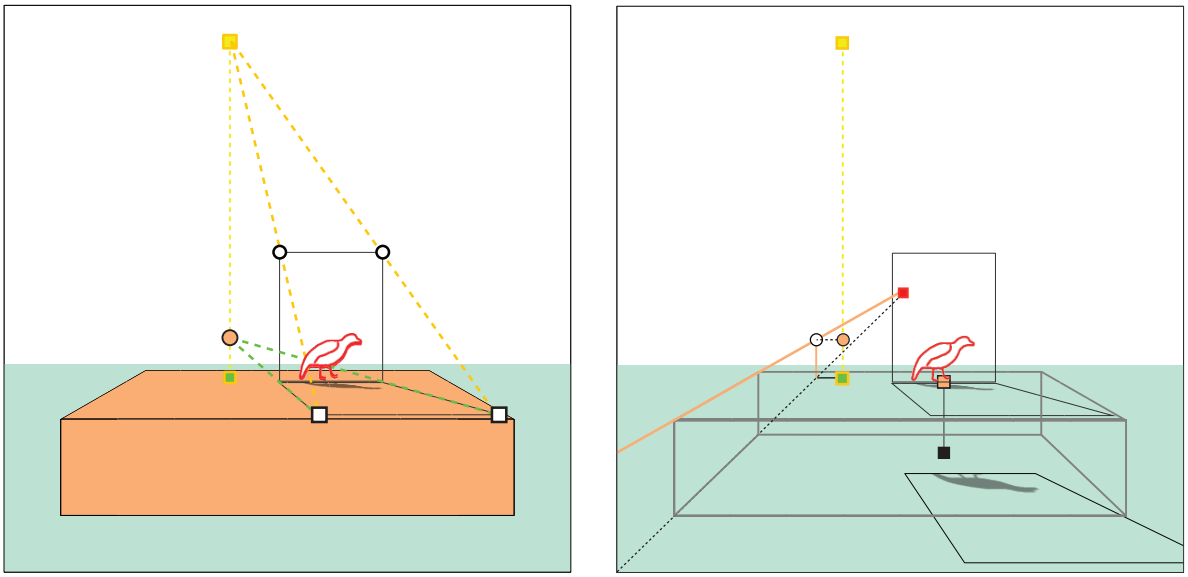


Figure 5.22: Panel shadow on an elevated ground. Ground rays and the light vanishing point on the elevated surface are used. Only the part of the panel above the height defined by the elevated plane need be considered.

Figure 5.23 (b) shows the construction to evaluate the elevated *lvp*. As in Section 5.4.1, the main orthogonal plane is the location from which the elevation is geometrically evaluated[15]. In addition, since the bird panel rests on the top of the box, the attachment point elevated to the top face intersects the panel base, coinciding with the panel origin, as shown in Figure 5.23 (b). The elevated reference points being known, the elevated ground rays with their corresponding light rays produce the boundary of the projected shadow on the top face of the box, as shown in Figure 5.23 (a). The panel mask is warped within this projected quadrilateral boundary.

In Figure 5.22, because the woman stands on the ground, the elevated attachment point is above the panel origin. This, the elevated ground rays do not pass through the bottom corners of the panel. In such a case, the content mask of the part of the panel below the elevated ground is clipped. The construction ignores the part of the panel between the floor and the elevated attachment point. In fact, only the transparent yellow region in Figure 5.22 is used in the construction of the elevated shadow.

[15]The naive elevation construction using the line from the *pvp* through the *lvp* runs the risk of using vertically aligned lines making the evaluation ill-defined as in Figure 5.6.



(a) (b)

Figure 5.23: Constructions for shadows on an elevated ground. (a) shows the construction that evaluates the projected shadow boundary mapping. (b) shows the construction to evaluate the elevated lvp: the main orthogonal plane is used to elevated the lvp by the height of the block.

On Vertical Planes. Some illumination configurations create visible shadows on planes parallel to the image plane. Two specific illumination positions often produce such shadows: a light close to the image plane, which has its lvp between the viewer and a scene object, and a pseudo-illuminant. behind the viewer. In both cases, a plane that directly faces the viewer, perpendicular to the floor and parallel to the image plane, such as a wall in the scene depth, might intercept a shadow extending in depth. Then the shadow that points away from the viewer continues on the vertical face where it is visible.

Figure 5.24 illustrates a typical scene where a shadow falls on a shadowed plane parallel to the image plane and is visible. The scene contains a panel object and a back wall illuminated by a pseudo-sun, so that shadows fall on both the ground plane and the back wall. To construct the projected shadow on the ground plane, light rays are drawn from the projected pseudo-sun location (below the lvp on the horizon) to the panel boundary corners and ground rays from the lvp to the attachment points of the panel vertical edges.

Where the ground shadow quadrilateral intersects the base of the wall, vertical lines are drawn. If they first intersect the light rays through the top corners of the panel boundary, the shadow is complete on the back wall as shown. Contrarily, if they first intersect the top of the

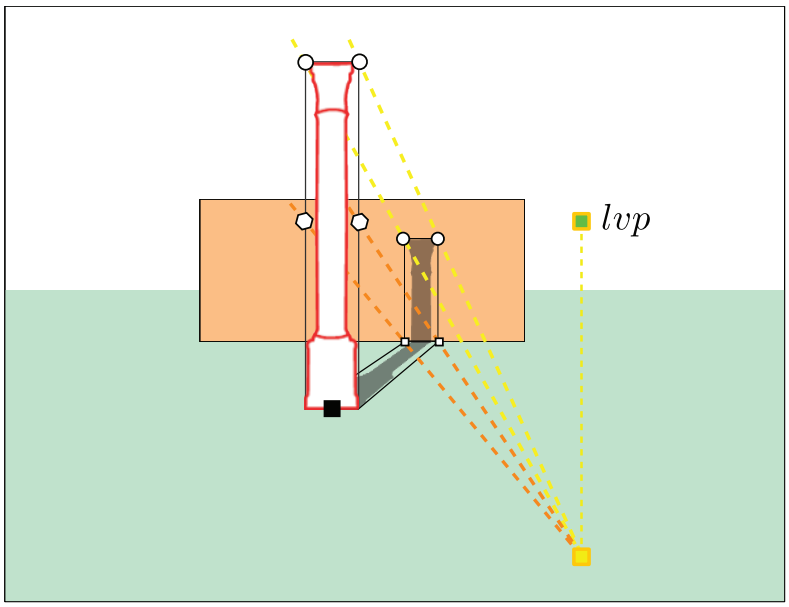
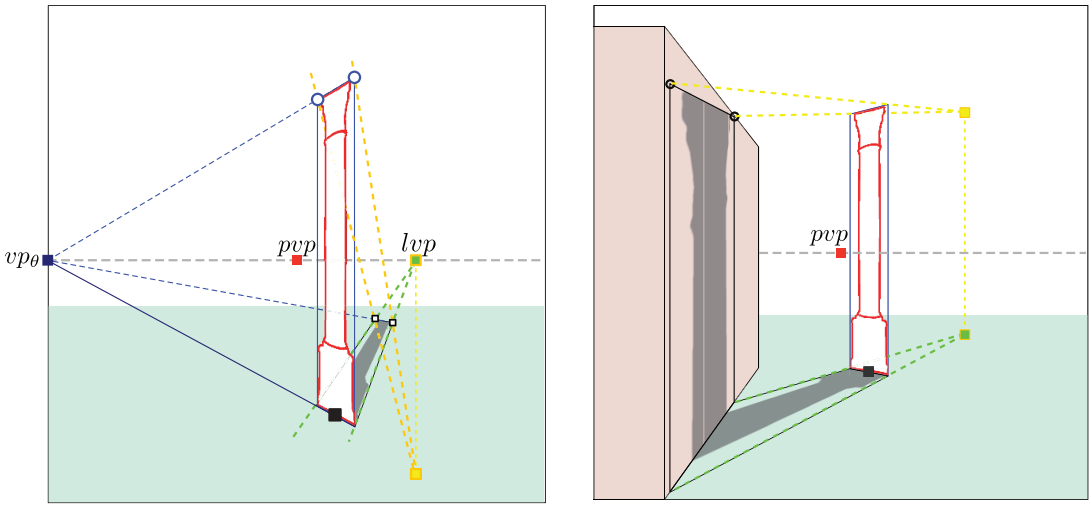


Figure 5.24: Pseudo-illuminant creating a ground shadow that continues onto a wall parallel to the image plane and is visible to the viewer. Light rays map the relevant regions.

shadowed surface then the shadow continues onto another, possibly invisible, elevated surface. Figure 5.24 shows the light rays that estimate the shadow projections of the panel boundary on the two shadowed planes. When all shadow border parts have been determined, the robust alternative is to project back the corners of each shadow quadrilateral to the panel with light rays, so as to subdivide the content used for each mapping. Another alternative is to keep the full shadow projected on the plane and to apply a clip when rendering the shadow: for the vertical wall, for example, the full shadow projected onto the shadowed plane, including below the ground, can be stored as long as it is clipped to the wall face at rendering time. This alternative saves computation should the wall height be modified.

Projecting a shadow onto a plane parallel to the shadowing panel is a simple scale. Thus, a simple scale applied to the image mask renders the shadow properly.

Rotated planes. Similar constructions apply to projected shadows where the shadowing and shadowed planes are in arbitrary orientations. For example, the shadowing plane might not be perpendicular to the line of sight. The ground shadow created by a rotated panel is described, then the construction of its shadow projected onto a plane perpendicular to both the floor and the image plane is briefly explained.



(a) (b)

Figure 5.25: Construction of projected shadow for rotated panels on the ground plane and on a lateral vertical plane. (a) The parallel lines converge at vp_θ on the horizon. (b) Some of the construction lines necessary to project the shadow of the rotated panel onto the medial face of a block. (For further information see Appendix C.3.)

The ground shadow of a vertically rotated panel is constructed by considering the ground rays that pass through the attachment points of the vertical edges of the panel boundary and the light rays that pass through the corners of the panel boundary. The construction that rotates the panel, Figure 5.16, directs the placement of the rays that evaluate its ground shadow. Figure 5.25 (a) shows for a panel vertically rotated by θ the construction of its ground shadow for a pseudo-illumination: the non-vertical edges of the panel boundary and their corresponding ground shadow edges intersect at vp_θ on the horizon.

The shadow on a ground plane intercepted by an arbitrary vertical plane is delimited by light rays to the intersection of the ground plane and the vertical surface. Figure 5.25 (b) shows such a shadow on the medial face of a block. The shadow on the face exists if the ground shadow quadrilateral intersects at least once the base edge of the block face: the two ground intersections with the line of this edge define two verticals. The panel boundary shadow on the plane of the medial face is bounded by the intersections of each vertical with a light ray through the panel boundary top corner.

The shadow construction for panels apply equally to the vertical faces of blocks, generating the volume of the projected shadow.

Projected shadows are perceptually important, as much for showing contours of the shadowed surface as for defining the properties of the shadowing object [14]. In paintings fuzzy edges usually deemphasize the shadows[16] because they are rarely significant to the picture's narrative. Here, the shadow constructions demonstrate that double projection is handled correctly by geometric algorithms. Shadows continuing across non-parallel surfaces, which reinforce the geometry of the shadowed and shadowing surfaces, are of primary interest.

5.6 Discussion of Construction

This chapter explained the construction line algorithms available in my framework. It may be hard to dissociate those that are direct applications of conventional perspective (Chapter 4) from the framework algorithms that require adaptation of perspective practice principles in a different context. The panel rotation is a particular affordance that I found useful for my framework because it takes advantage of the computational environment to adjust a local representation by modifying its individual perspective. I made that construction and turned it into an algorithm. By extension, I had to derive its projected shadows, in particular the construction to project a rotated shadow onto a vertical wall along a tile orthogonal is non-trivial: the construction of Figure 5.25 (b) is explained further and analyzed for correctness in Appendix C.3.

Indeed, automating artists' construction practices is seldom a straightforward application owing to the tacit knowledge encompassed in artists' methods. For example, while manuals prescribe the usage of distinct light vanishing points for different shadowed surfaces [17], they do not mention the need of an object attachment point or of its elevated version: artists know where to place the lines in relation to the shadowing and shadowed surfaces. Furthermore, clipping to discard the invisible part that projects through infinity is necessary to make the algorithm stable and correct. Artists handle those cases intuitively using their tacit knowledge: no explanations are required, nor is the problem made explicit, it simply does not exist. My implementation, which interactively changes object and light placements, made the problem appear and the clipping solution was found.

My framework promotes exploring of the geometry of a perspective picture: the construction lines displayed when light points and objects are moved link the appearance to the geometry that produces it. My framework helped me to investigate, to understand and to generalize perspective constructed on the canvas surface. In fact, some capabilities not described in this

[16] My framework implementation has attributes for shadow appearance: colour, opacity and blurriness (§ D.4.2). Figure 5.19 illustrates fuzzy shadow edges, and the fuzziness increases with the distance to the lvp.

chapter were added in my framework as a response to my exploration. Investigating the unsatisfying appearance of the projected shadow of a panel for some light placements, I realized that the effect could be improved by using a rotated version of the panel to cast the shadow. In those circumstances, the shadowing surface is invisible, but is used for computing the shadow of an visible panel, which is not rotated. Furthermore, I experimented with the idea of multiple perspectives for the geometry of blocks. Block constructions need not adhere to the principal perspective of the picture. My framework allows its user to move a block's vanishing point to explore representations whose geometry is abstract. Those features are not fully developed but the exploration described here is included in Chapter 8.

5.7 Conclusion

This chapter described my framework, which has the unique feature of incorporating into a picture objects that exist in a simulated 3D environment. It is possible because the objects are injected into a 2D environment with pre-defined view parameters. Each object is linked to the perspective by its attachment point. Thus, it can be manipulated predictably.

Geometric algorithms compute object appearance without reference to 3D. They modify manipulated objects so as to maintain coherence with the simulated 3D environment. The development of the geometric algorithms had several objectives:

- to demonstrate that artists do not implicitly introduce 3D analysis into their constructions on the canvas[17],
- to show by example that complete calculations of perspective and lighting can be performed using only 2D image plane coordinates,
- to demonstrate that the tacit knowledge of artists can be be made explicit so that the constructions can be formalized, and
- to provide a reference implementation to which the research described in Chapter 6 can be compared.

Instantiating construction lines for simulating 3D using geometric calculation is tedious, although it certainly improves one's geometric intuition. Thus, the next chapter reformulates those constructions in terms of 2D projective geometry. An algebraic formalism systematically

[17]Only the projection invariances of straightness and incidence, and the nature of parallel lines are necessary.

computes the geometric construction. Its condensed computations are easier to implement, and are well-adapted to current graphics hardware.

Artists' constructions illustrate the salience of the attachment point for human vision. In a world with gravity the feet of most figures coincide with the attachment point, rooting them firmly on the ground plane. Unnoticed properties of the ground—solidity, opacity—guarantee this relationship with the point of contact that is so central in human perception of depth. Consider objects that are dissociated from their attachment points, such as flying birds, hot air balloons or clouds. This is a set of objects for which it is notoriously difficult to estimate depth. Thus, treating apprehension of location in space as a ground location, plus a height above it gives many advantages. For example, seeing a person walking at a distance, one does not perceive an (x, y, z) location of the person's shoulder, but rather how far away the person is on the ground plane, and how tall[18]. My framework, by demanding its user to specify the attachment points, takes advantage of the human operational approach to space, dissociating the ground plane from the height above it.

[18]In human vision it is a relation on the ground surface with size constancy above its depth link that causes one to estimate effortlessly the person's height.

Chapter 6

Formalism

> *The interplay between the intuitive geometric picture (which tells you what you should be expecting), and the precise formulation in terms of coordinates (which allows you to cash in on your intuition) is a fascinating aspect of algebraic geometry.*
>
> Miles Reid (Undergraduate Algebraic Geometry [91, p. 9])

This chapter presents a 2D mathematical formalism that reproduces in algebra the geometric constructions described in Chapter 5. I convert the construction lines of artists into lines in $\mathbb{RP}^2$, the real projective plane. Because the algebra reproduces the geometry of artists it is 2D. Non-singular 3×3 matrices, up to a constant, form a group describing projective mappings in $\mathbb{RP}^2$, without leaving the projective plane.

There are four reasons to investigate this formalization. Complementing geometric intuition, algebra provides a more secure foundation and permits useful computation. Algebra also shows a different side of the construction. Of particular interest is the sufficiency of the attachment point concept, and its interaction with height. The group composition operator, 2D projective matrix multiplication, makes complex transformations out of simple ones. Lastly, it links the 2D geometry to the 3D algebra of computer graphics.

Furthermore on the practical side, the algebraic formulation eases the implementation of the framework. In particular, it can use contemporary graphics hardware.

This chapter describes the environment in which the 2D projective mappings function: the Renaissance canvas space, $\mathbb{RCS}$. Then, the fundamental matrices to evaluate the two types of perspective points are derived: first for points on the ground plane, later generalized for general points, above the ground plane points. Appendix A provides more matrices that complete the

manipulations of Chapter 5. Appendix B shows that the ground plane mappings match the optical physics of 3D scene geometry. Appendix C further grounds the formalism in general projective geometry using the center of projection that exists on the ℝ℃𝕊.

6.1 Renaissance Canvas Space: ℝ℃𝕊

The formalism is not an ad-hoc collection of mappings that happens to create pictures in perspective, but a well-defined set of mappings corresponding to the constructions of artists. They exist in a space of points and lines, in which two points define a line, and two lines a point, which is at infinity if the lines are parallel. This describes $\mathbb{RP}^2$, the projective plane. The Renaissance artist has, in addition, pre-chosen the view-parameters, which determines the exact form of the transformations. Thus, this constrained $\mathbb{RP}^2$ is given its own name, **Renaissance canvas space**, ℝ℃𝕊.

Piero's diagonal and the distance point constructions, which are specific to the view parameters, place parts of the 3D physical space on the ℝ℃𝕊.

- The 3D ground half-plane, Piero's physical floor, is on the ℝ℃𝕊 below the baseline.
- The eye point is on the ℝ℃𝕊 at the distance point.

It remains to determine the ℝ℃𝕊 location of the above ground part of the 3D physical quarter volume behind the image plane. This physical space behind the image plane is best treated as a set of horizontal half-planes. It is then natural to transfer each 3D physical **slice**, a horizontal half-plane, to the ℝ℃𝕊 by similarly rotating it down 90° around its **joint line**, which is its intersection with the ℝ℃𝕊. Doing so is in accord with the construction in Figure 5.5.

6.1.1 Physical and Perspective Spaces of the ℝ℃𝕊

Figure 6.1 illustrates the ℝ℃𝕊 with its two sets of elements.

- The **perspective** ℝ℃𝕊 (black) consists of
  1. the dp and the pvp representing the eye point,
  2. the perspective ground plane, and
  3. the slices parallel to it: the horizontal half-planes above the ground plane.

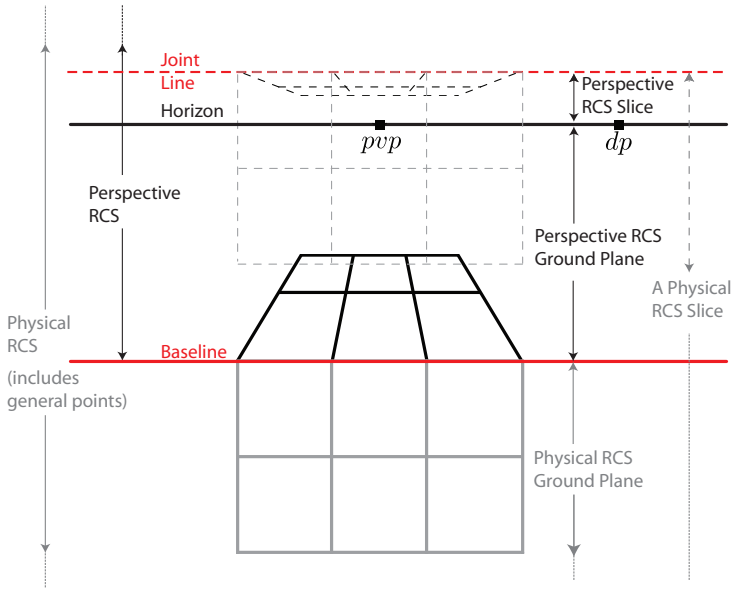


Figure 6.1: The formalism maps on canvas points from the ℝℂ𝕊 physical space (gray) to the ℝℂ𝕊 perspective picture (black), the baseline being common. Similarly, the joint lines, lines parallel to the baseline have ℝℂ𝕊 perspective lines that correspond to their ℝℂ𝕊 physical and 3D physical space equivalents. Shown above is the ground plane, of which the joint line is the baseline, and an elevated slice.

Figure 6.1 shows the pvp and dp, the perspective ground plane (the 3×2 tiled floor); and an elevated slice (the dashed 3×2 perspective grid).

- The **physical** ℝℂ𝕊 (gray) consists of slices above the ground plane and behind the canvas: each slice is drawn downward on the ℝℂ𝕊. Because the slices overlap, it is rare to draw several on a single diagram. The slice at the baseline is the physical ground plane.

  Figure 6.1 shows the physical ground plane (the 3×2 grid) and a parallel slice above it (the dashed 3×2 grid), which corresponds to the elevated perspective grid drawn.

6.1.2 The Attachment Point

There is a problem in the ℝℂ𝕊: overlaps exist, many of them, as everything—ground planes and slices—is included on the same surface, the image plane. For example, in the perspective ℝℂ𝕊 other points not at infinity are drawn on the horizon. A joint line and its entire perspective slice coincide with the horizon. Artists easily differentiate overlapping points by eye, moving picture elements on the canvas if necessary.

By using only the physical representation for the perspective ground plane, Piero could place the physical and perspective $\mathbb{RCS}$ without overlap avoiding the problem. By singling out the two ground planes he showed something important: the ground plane is special. It allow us to define uniquely any point within the perspective or the physical $\mathbb{RCS}$.

The solution is, thus, the attachment point. This point on the ground plane is the fundamental mechanism to work directly on the image plane: it differentiates collocated locations on the canvas.

6.1.3 Homogeneous Coordinates

To derive the $\mathbb{RCS}$ transformations, the 2D projective algebraic representations are homogeneous coordinates. My notation follows Davis', who recommended homogeneous coordinates for practical computation in computer graphics [23, p. 9–14]: a 2D point is a column triplet; its dual line a row triplet. Specifically, the essential properties of the real projective plane $\mathbb{RP}^2$ used for my derivations are as follows.

"Every vector of three real numbers, $(x, y, w)^T$, where at least one of the numbers is non-zero, corresponds to a point in two-dimensional projective geometry. The coordinates for a point are not unique; if α is any non-zero real number, then the coordinates $(x, y, w)^T$ and $(\alpha x, \alpha y, \alpha w)^T$ correspond to exactly the same point.

If the w-coordinate is non-zero, it will correspond to a Euclidean point, but if $w = 0$ (and at least one of x or y is non-zero), it will correspond to a "point at infinity" [...].

Furthermore, the allowable transformations in (two-dimensional) projective geometry correspond to multiplication by arbitrary non-singular 3×3 matrices. Obviously, if two matrices are related by the fact that one is a constant non-zero multiple of the other, they represent the same transformation." [23, p. 9]

I use the swung dash $\sim$ notation, as is common in projective geometry, [91, p. 11], to mean 'is the same point as': $(x, y, w)^T \sim \alpha(x, y, w)^T$ if $\alpha \in \mathbb{R}\backslash\{0\}$.

Next the fundamental $\mathbb{RCS}$ transformations are derived. They define a mapping between perspective $\mathbb{RCS}$ points and physical $\mathbb{RCS}$ points parametrized by the view parameters, which are the perspective $\mathbb{RCS}$ locations of the pvp and dp. First the transformations between points on the two ground planes are derived, then the transformations are generalized to points that are above the ground planes.

6.2 Ground Plane Points

Following Piero's diagram (Figure 4.12), Section 5.3 describes the construction that maps a point on the physical ground plane to the corresponding point on the perspective ground plane. It uses six lines intersecting the baseline and Piero's diagonals to determine its perspective location. My formalism instead deduces the perspective-dependent non-singular 3×3 matrix that maps a ground point of the physical ℝℂS to its perspective equivalent.

In effect, there exists a 2D projective transformation that maps between the physical ℝℂS ground plane and the perspective one. The mapping is invertible. Its inverse transforms a ground point of the perspective ℝℂS to its physical equivalent: the inverted geometric construction discussed is in Section 5.3.

6.2.1 Description

The tiled floor on the physical ℝℂS ground plane is a rectangular area containing square tiles of side s. It defines two sets of parallel lines at right angles that are perpendicular and parallel to the image plane. The tiled floor on the perspective ℝℂS ground plane is the perspective image of these orthogonal and horizontal lines, defined completely by the viewpoint, the distance point, and the size of the tiles, measured at the baseline.

Figure 6.2 shows a single row of perspective tiles with the corresponding physical ones directly below. The 2D coordinate system for the ℝℂS has its origin, **O**, at the center of the baseline; x increases to the right, y upward. The principal vanishing point of the perspective picture, abbreviated as pvp, labelled **V** in mathematical expressions[1], is on the horizon directly above the origin **O**. **V** is at a height $\mathbf{E}_h$ above the baseline: the height of the viewer's eye. The orthogonals of the perspective floor converge to **V**. The distance point, dp, labelled **D** in mathematical expressions, is on the horizon, separated from **V** by $\mathbf{E}_d$: the distance of the viewer from the image plane. The foreshortening of the perspective floor depends of **D**.

6.2.2 Derivation

With the coordinate system centered at **O**, the goal is to determine the matrix $\mathcal{Q}$ that maps a point on the physical ground plane (gray in Figure 6.2) to its location on the perspective

[1]The ℝℂS coordinates of **V** and **D** are expressed in terms of $\mathbf{E}_d$ and $\mathbf{E}_h$, which depend on the 3D viewer's location; the derivations match their optical equivalents (Appendix B). While pvp and dp are appropriate abbreviations for text use; the single letter **D** and **V** respects the notation used for points in the mathematical expressions.

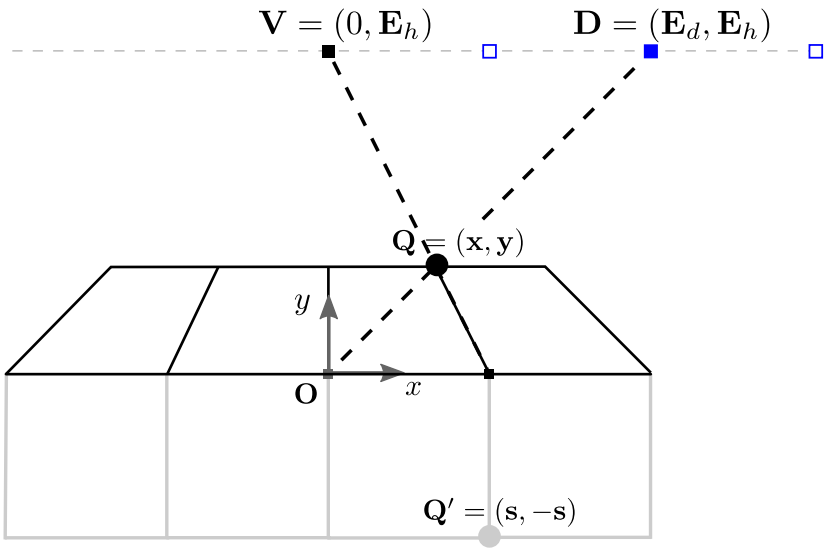


Figure 6.2: Geometric positions on the $\mathbb{RCS}$ between the perspective tiled floor and the physical one beneath it. The mapping generalizes the construction used to position the tile corners of the two spaces. The shared baseline remains unchanged. The perspective tiled floor appearance depends of the distance point, **D**: the square filled in blue. (The two other outlined squares are alternate positions for **D** used in my framework.)

ground plane (black). ω is determined by the mappings between four unique, non-collinear ground points, each a correspondence between a physical ground point below the baseline and its perspective equivalent above it. All points on the baseline, $(\mathbf{x}_b, 0, 1)^T$ are common. On the physical $\mathbb{RCS}$ ground plane the point at infinity $(0, -\infty, 1)^T \sim (0, -1, 0)^T$, an ideal point in $\mathbb{RP}^2$, corresponds to the principal vanishing point on the perspective $\mathbb{RCS}$ ground plane, $\mathbf{V} = (0, \mathbf{E}_h, 1)^T$, an actual point on the perspective picture. Similarly[2], on the physical $\mathbb{RCS}$ ground plane the point at infinity $(\infty, -\infty, 1)^T \sim (1, -1, 0)^T$ corresponds to the distance point on the perspective $\mathbb{RCS}$ ground plane, $\mathbf{D} = (\mathbf{E}_d, \mathbf{E}_h, 1)^T$. In effect, using the following three[3]

[2]This mapping, which also uses an ideal point, demands projective geometry intuition. An alternative is to use the mapping between $\mathbf{Q}' = (s, -s, 1)^T$ and $\mathbf{Q} = (s\mathbf{E}_d, s\mathbf{E}_h, s + \mathbf{E}_d)^T$, which are shown in Figure 6.2.

[3]The first mapping holds for all **x** coordinates on the baseline thus setting more than one point.

mappings,

$$\begin{pmatrix}\mathbf{x}_b\\0\\1\end{pmatrix}\mapsto\begin{pmatrix}\mathbf{x}_b\\0\\1\end{pmatrix},\quad\begin{pmatrix}0\\-1\\0\end{pmatrix}\mapsto\mathbf{V}=\begin{pmatrix}0\\\mathbf{E}_h\\1\end{pmatrix},\quad\text{and}\quad\begin{pmatrix}1\\-1\\0\end{pmatrix}\mapsto\mathbf{D}=\begin{pmatrix}\mathbf{E}_d\\\mathbf{E}_h\\1\end{pmatrix},$$

the matrix ϖ is determined:

$$\begin{pmatrix}\mathbf{Q}_x\\\mathbf{Q}_y\\1\end{pmatrix}\sim\varpi\begin{pmatrix}\mathbf{Q}'_x\\\mathbf{Q}'_y\\1\end{pmatrix}\sim\begin{pmatrix}-\mathbf{E}_d&0&0\\0&\mathbf{E}_h&0\\0&1&-\mathbf{E}_d\end{pmatrix}\begin{pmatrix}\mathbf{Q}'_x\\\mathbf{Q}'_y\\1\end{pmatrix}. \tag{6.1}$$

Reciprocally, the inverse matrix ϖ^{-1}, which transforms a $\mathbb{RCS}$ perspective floor point, $\mathbf{Q}$, to a $\mathbb{RCS}$ physical floor point, $\mathbf{Q}'$ is

$$\begin{pmatrix}\mathbf{Q}'_x\\\mathbf{Q}'_y\\1\end{pmatrix}\sim\varpi^{-1}\begin{pmatrix}\mathbf{Q}_x\\\mathbf{Q}_y\\1\end{pmatrix}\sim\begin{pmatrix}-\mathbf{E}_h&0&0\\0&\mathbf{E}_d&0\\0&1&-\mathbf{E}_h\end{pmatrix}\begin{pmatrix}\mathbf{Q}_x\\\mathbf{Q}_y\\1\end{pmatrix}. \tag{6.2}$$

ϖ^{-1} has an interesting duality with ϖ: the coordinates $\mathbf{E}_d$ and $\mathbf{E}_h$ are swapped in the entries of the matrices, indicating that there is an 'inverse' perspective with x and z axes interchanged. Appendix C discusses the underlying 3D geometry; Figure C.4 shows the coordinates swapped and their relationship to the auxiliary eye point. The duality of ϖ and ϖ^{-1} in terms of points and lines in projective geometry is explained in Section A.2.

6.3 General Points

With the mappings ϖ and ϖ^{-1}, which move points between the $\mathbb{RCS}$ physical and perspective ground planes, we are now ready to derive mappings for points elevated above the ground plane. To render blocks and panels in perspective, ϖ and ϖ^{-1} generalizations are required.

6.3.1 Description

A general point belongs to a slice, which is parallel to the ground plane (§ 6.1): it is defined by its position on the ground plane, which is its attachment point, and its physical height $\mathbf{c}'$. The attachment point is a 2D point $\mathbf{Q}$ on the perspective floor which corresponds to the 2D point $\mathbf{Q}'$ on the physical ground plane. On the perspective $\mathbb{RCS}$, the general point, $\mathbf{P}$ is $\mathbf{c}'$ above $\mathbf{Q}$ only when $\mathbf{Q}$ is on the baseline, whereas on the physical $\mathbb{RCS}$, $\mathbf{P}'$ is always $\mathbf{c}'$ above $\mathbf{Q}'$.

There are at least two[4] different constructions to derive the perspective ℝℂS location of a general point from its definition. The derivation presented here is based on the artist's intersection construction for evaluating a foreshortened height (Figure 5.5). Indeed, the derivation, following Figure 6.3, uses the tiled floor on slice $\mathbf{c}'$ to evaluate a general point's position on the perspective picture.

6.3.2 Derivation

The general 2D projective transformation, which maps a point at height $\mathbf{c}'$ above its attachment point, $\mathbf{Q}'$ on the physical ℝℂS floor, to its location, $\mathbf{P}$, above the floor on the perspective ℝℂS, is derived in what follows. Figure 6.3 shows the physical ℝℂS slice $\mathbf{c}'$ in relation to its ℝℂS perspective equivalent: the principal vanishing and distance points on the horizon and the tile corners on the slice joint line are sufficient to deduce the foreshortening of the elevated tiles.

A translation upward is needed first, positioning the construction on the $\mathbf{c}'$ slice: its joint line is the location where the physical ℝℂS slice corresponds to its perspective ℝℂS equivalent. Once positioned at the joint line, a generalized mapping of ϖ is needed to consider the different heights separating the horizon from the joint line used for each given slice: $\varpi_{\mathbf{c}'}$ is the generalized mapping. But some pre- and post-considerations with respect to the invariance line connecting both spaces apply: a change of basis is required so that invariance holds for the joint line of the slice, rather than for the baseline. Indeed, ϖ keeps the points on the baseline, $(\mathbf{x}_b, 0, 1)^T$, fixed. To keep the joint line points, $(\mathbf{x}_b, \mathbf{c}', 1)^T$, invariant, a change of basis is needed. The matrix $\phi_{\mathbf{c}'}$ that maps points on the physical ℝℂS slice $\mathbf{c}'$ to points on its perspective ℝℂS representation demands a change of basis back, $\tau_{\mathbf{c}'}^{-1}$, and forth, $\tau_{\mathbf{c}'}$, around $\varpi_{\mathbf{c}'}$.

$$\phi_{\mathbf{c}'} = \tau_{\mathbf{c}'}\varpi_{\mathbf{c}'}\tau_{\mathbf{c}'}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \mathbf{c}' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h - \mathbf{c}' & 0 \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\mathbf{c}' \\ 0 & 0 & 1 \end{pmatrix}$$

$$\phi_{\mathbf{c}'} = \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h & -\mathbf{c}'(\mathbf{E}_h + \mathbf{E}_d) \\ 0 & 1 & -(\mathbf{c}' + \mathbf{E}_d) \end{pmatrix}.$$

[4] An alternate derivation is described in Appendix A.1: it reveals a new geometric construction, building in the perspective space a distinct ground with its tiled floor. It results in the same matrix and does not includes a change of basis, as the derivation here requires.

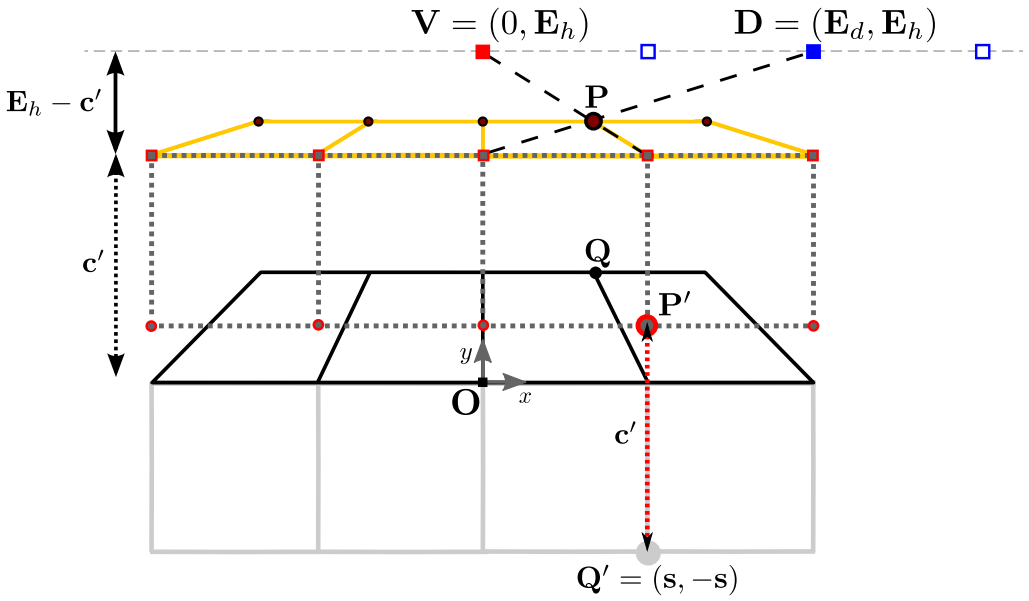


Figure 6.3: The general point evaluation, $\mathbf{P}$, is directly done on the slice $\mathbf{c}'$ above the ground plane. The physical $\mathbb{RCS}$ slice (dotted grey), elevated from the physical floor by $\mathbf{c}'$, is used. Its joint line acts as the baseline, to construct its perspective $\mathbb{RCS}$ equivalent (yellow).

The mapping $\phi_{\mathbf{c}'}$ transforms $\mathbf{P}'$ to $\mathbf{P}$ thus to derive the transformation, $\pi_{\mathbf{c}'}$ that applies to $\mathbf{Q}'$ the initial translation up to the physical $\mathbb{RCS}$ slice $\mathbf{c}'$ is added. Indeed, we have

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \pi_{\mathbf{c}'} \begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix} \sim \phi_{\mathbf{c}'} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \mathbf{c}' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix},$$

which reduces to

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h & -\mathbf{c}'\mathbf{E}_d \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix}. \tag{6.3}$$

$\pi_{\mathbf{c}'}$ computes the $\mathbb{RCS}$ perspective position of the general point $\mathbf{P}$ given its physical $\mathbb{RCS}$ attachment point $\mathbf{Q}'$ and physical height $\mathbf{c}'$. Note that $\pi_{\mathbf{c}'} = \varpi$ for points directly on the floor, which have height $\mathbf{c}' = 0$. It is interesting to note that the construction and its corresponding projective transformation, π, work for points that are above the horizon: $\mathbf{c}' > \mathbf{E}_h$. The joint line is then above the horizon and this flipped position creates an elevated tiling that in perspective converges to $\mathbf{V}$ from above rather than below.

By extension on the perspective $\mathbb{RCS}$, a direct transformation maps from an attachment point, $\mathbf{Q}$, and a height $\mathbf{c}'$, to the location $\mathbf{P}$:

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \Pi_{\mathbf{c}'} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \sim \pi_{\mathbf{c}'} \varpi^{-1} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & \dfrac{\mathbf{E}_h - \mathbf{c}'}{\mathbf{E}_h} & \mathbf{c}' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix}. \tag{6.4}$$

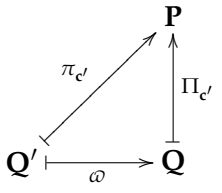
Figure 6.3 shows that the general point mapping, $\Pi_{\mathbf{c}'}$, naturally builds the structure for blocks in perspective, which are the geometric shapes that approximate the scene architecture in the framework. The $\mathbb{RCS}$ and its transformations are also well-adapted for working with cartoons and projected shadows. A variety of transformations comparable to $\Pi_{\mathbf{c}'}$ implemented for my framework to sustain user manipulations are derived in Appendix A. Those functioning within the perspective $\mathbb{RCS}$ are denoted with upper case Greek letters: $\Theta_{\mathbf{c}'}$ for rotation, $\Gamma_{\mathbf{c}'}$ for pushing in depth and $\Psi_{\mathbf{c}'}$ for projected shadow on slices.

6.4 Summary

This chapter presented derivations of the 2D projective mappings needed to create the geometry and to manipulate the elements of a perspective picture. For a physical ground plane location, $\mathbf{Q}'$, and a height 2D homogeneous matrix multiplication calculates the corresponding location on the perspective $\mathbb{RCS}$ (Equation 6.3). Alternatively, within the perspective $\mathbb{RCS}$ itself, the perspective coordinates of any point can be evaluated (Equation 6.4).

This second type of mapping, $\Pi_{\mathbf{c}'}$, (Equation 6.4) in more useful because it maps the perspective $\mathbb{RCS}$ to itself: its independence of $\mathbf{E}_d$ is most surprising. $\mathbf{E}_d$ only fixes the location of $\mathbf{Q}$ on the perspective $\mathbb{RCS}$ ground plane and nothing further. The foreshortened height of the point is solely a function of $\mathbf{Q}$ and the proportional distance to the horizon. The most striking observation of the derived mappings is the interdependence of the entities of the $\mathbb{RCS}$. In particular, we used the following mappings between points on the physical and perspective $\mathbb{RCS}$ ground and perspective $\mathbb{RCS}$ slices:

where ϖ depends on $\mathbf{E}_h$ and $\mathbf{E}_d$, $\pi_{\mathbf{c}'}$ on $\mathbf{E}_h, \mathbf{E}_d$ and $\mathbf{c}'$ and $\Pi_{\mathbf{c}'}$ on $\mathbf{E}_h$ and $\mathbf{c}'$.



In building the formalism the focus has been on the practical calculations involving elements that are visible, omitting their effects on the scene, which is implicitly clipped, the image plane functioning like the near plane. But its projective mappings have details that are mathematically interesting, as relations outside the visible space behind the image plane exist. In particular, the 3D physical space in front of the image plane is included by a center of projection on the ℝ℃𝕊: Appendix C describes this ℝ℃𝕊 invariant point, which is a ℝ℃𝕊 center corresponding to the projection of the 3D eye, and is distinct from the *pvp*. In terms of projective geometry those theoretical details are interesting, even if they have no practical purpose for this thesis, which is the application of 2D projective geometry to computer graphics.

6.5 Discussion

This chapter described the Renaissance canvas space, ℝ℃𝕊, and its fundamental matrices to evaluate general points in the perspective picture. Algebraic computations for the other constructions presented in Chapter 5 are described in Appendix A. In particular, the mappings to compute a rotated panel and projected shadows are derived. The transformation for shadows on the ground plane is generalized to any horizontal shadowed planes, the ℝ℃𝕊 slices.

As the epigraph to this chapter suggests the algebraic formalization of geometry presented in this chapter and in Appendix A shows aspects of geometric objects that are not obvious when thinking geometrically. For example, in Chapter 5 the attachment point may have seemed an appendage to the point itself, probably secondary while the point is primary. But when the construction is put into algebra it becomes clear that the transformation between $\mathbf{Q}'$ and $\mathbf{Q}$ is the central idea. Its transformation ϖ (Equation 6.1) derived in this chapter, is used throughout the formalization of the transformations that manipulate panels (§ A.3). In fact, one comes to think of the ground point transformation as moving a vertical line from one space to another, where the physical line is parametrized by $\mathbf{c}'$ and the perspective line by $\mathbf{c}$. (The two heights have a simple relation: $\mathbf{c} = (1 - \mathbf{Q}_y/\mathbf{E}_h)\mathbf{c}'$, as derived from $\Pi_{\mathbf{c}'}$.)

Artists by developing their constructions have isolated the lines within the context of the

task at hand, which have a concrete (human rational) meaning and an influence for developing the geometry of a perspective picture. Points on the ground are commonly tacit knowledge in their practice, but artists do refer to the light vanishing point, which is required to depict projected shadows. Making those points implicit with the attachment point representation, is the mechanism that encodes just enough 3D information accessible in 2D: the attachment points enable to calculate the perspective picture from the canvas surface. This observation is an important contribution of this thesis. The formalization of projected shadows (§ A.4) is only possible due to attachment points of both the shadowing and shadowed surfaces.

In addition, turning artists' constructive geometry into algebra produces mappings that are interesting, because traceable: with the ℝℂ𝕊 one can see the geometry in the algebra it produces. Indeed, working on the algebra expands the constructive concept: a new construction was revealed (§ A.1). By following both the practical geometry and the mathematical formalism we understand each better and see the principles of portraying 3D on a flat image plane in a new light.

In particular, this chapter has shown that there are one-to-one mappings between points on the perspective ℝℂ𝕊, labelled by attachment point and height $\mathbf{Q}, \mathbf{c}'$ and points in the physical ℝℂ𝕊 labelled as $\mathbf{Q}', \mathbf{c}'$, the two mappings $\pi_{\mathbf{c}'}$ and $\pi_{\mathbf{c}'}^{-1}$.[5] Isomorphism is a corollary of the existence of the two mappings. Any linear operation, Ω', that can be performed in the physical ℝℂ𝕊 can be performed in the perspective ℝℂ𝕊 by the forming operator

$$\Omega \sim \pi_{\mathbf{c}'} \Omega' \pi_{\mathbf{c}'}^{-1}$$

and *vice versa*. This operation is revealed by applying the matrices of this chapter to derive the transformations that manipulate my framework elements in Appendix A. The operation Ω underlies the derivation of $\Theta_{\mathbf{c}'}$, which evaluates general points on a vertically rotated panel (§ A.3.3).

This isomorphism should not be unexpected. The 3×3 homogeneous matrix is parametrized by the height of the point above its attachment point. The parametrization may be the oddest appearing feature of my formalization. Yet putting scene features into transformation matrices is not unusual. View parameters are essential for creating the matrix that transforms the view frustum into a cube. Indeed, finding the intersection between a ray and a plane can be performed by matrix multiplication; if there is one plane and many rays, the plane parametrizes the matrix and the rays are arguments; and *vice versa*. But I have gone further breaking what seems to be an atomic concept, a point, into parts. The decomposition tells us something

[5] I write $\pi_{\mathbf{c}'}^{-1}$ rather than π_{c}^{-1} because artists seem to think of figures in terms of their height at the picture plane.

important that the Renaissance painters knew and that we should make more explicit in computer graphics: although we live in a 3D universe almost all our motion takes place in 2D; when we ask where something is we almost always get a 2D answer, with a ground place implicit in the answer. Thus, the separation implied by the constructions of Renaissance artists is understandable in terms of practical psychology in a world with universal gravity.

As a result of the Ω operator, the perspective picture gains its self-sufficient coherence, with the formalism resulting matrices: $\Theta_{\mathbf{c}'}$, $\Gamma_{\mathbf{c}'}$, $\Psi_{\mathbf{c}'}$ function directly on the perspective $\mathbb{RCS}$. My geometric construction of Chapter 5 that needs the rotated line on the physical $\mathbb{RCS}$, is replaced by $\Theta_{\mathbf{c}'}$ where the physical $\mathbb{RCS}$ is by-passed. Similarly to artists who once have learned the basis of perspective, its principles, develop the perspective of the picture directly. My formalism results enable to become free of the physical ground plane. The physical tiled floor is in fact a tool for the novice who needs to understand the geometry in play.

Chapter 7

Rendering

Computers are good at following instructions, but not at reading [an artist's] mind.
D.E. Knuth (The TEXbook, p. 9)

Chapters 5 and 6 focused on the description of the geometry computed by my framework to assist a user creating the spatial organization of a perspective picture. This chapter addresses the rendering in the perspective $\mathbb{RCS}$. It explains the mechanisms and the theory behind the painter's algorithm I developed and that my framework executes. My painter's algorithm resolves the occlusion between scene elements by ordering in the perspective $\mathbb{RCS}$.

The main difference between my painter's algorithm and the traditional painter's algorithm is that its order is scene-independent. My painter's algorithm is based on the 2D tiled floor structure, inducing a scene ordering that does not require individual polygons to be sequenced. In effect, the visibility problem is approached in a way similar to how artists solve it in practice, almost at the granularity of scene object. My framework does not rely on the depth buffer commonly used in computer graphics.

7.1 General

This section gives essential details to understand the method of rendering in my framework. To complete the framework descriptions Appendix D gathers implementation details, describing the prototype interface as a solution to permit the user to create a schematic representation of a scene and the data structures needed to store the full composition.

7.1.1 Program

My framework, whose graphical user interface is shown in Figure 5.1, is implemented in Java, using the Java2D API. For both the geometric construction method and the 2D collineation formalism I used Java2D geometric library classes, particularly the 2D point, 2D affine matrix and transformation classes. In addition, a class implementing the perspective transform as a projective 3×3 matrix is used to map the content in a rectilinear shape to its 2D perspective transformed form. No 3D computation is performed.

My framework demonstrates that the geometric approach presented in this thesis is applicable in practice. Its performance is not optimized. The prototype runs at interactive frame rates on a 3 GHz Pentium 4 Linux box with 1 Gbyte of memory provided only that perspective mappings of panel content get lazy evaluation. Shadow effects exert the heaviest computation load. As the number of panels with shadows falling on many surfaces grows, the interactivity deteriorates when the light is moved: lazy shadow evaluation is then necessary.

7.1.2 Elements

The elements I render are

- the background,
- the ground,
- the volumes, i.e., the blocks,
- the panels, each representing a 3D local object, and
- the shadows.

While the first two, which are strictly behind the others, are drawn first (background then ground), the others require careful visibility assessment. The creation of a fixed drawing sequence that provides correct occlusion is the goal of my rendering algorithm. Solving occlusions among volumes is the hard part. Having established the volume drawing sequence, it is then straightforward to intermix panels and shadows among them.

Rather than considering the 3D position of the eye with respect to the scene elements, my oblique approach uses canvas space patterns to determine visibility relations. I show that the relationship of scene elements to the position of the principal vanishing point, pvp, are almost sufficient to compute visibility[1]. Both drawing requirement and sequencing of the faces that

[1]Considering the dp position for some more complex volumes is necessary as the eye distance interacts with the tangential visibility of some curved elements.

make the elements of the picture can be calculated from their locations in the perspective ℝℂS. This paradigm is fundamentally different to 3D visibility calculations because it processes by ordering in 2D and not 3D. Such a paradigm is adapted to my framework as its purpose is to mimic traditional artistic practices, and artists essential work on canvas appearance to determine visibility. The positions of construction lines and existing volume faces are fundamental guides to artists when they revise and complete the 3D simulation of a scene from the flat surface standpoint. Following artistic practice, in my framework, the appearance of the scene visibility is managed by analyzing interactions among the picture elements.

7.2 Current Methods for Calculating Occlusion

In drawing a realistic image, viewpoint-relative visibility calculations are performed to determine occlusion within the image content. These calculations, and the input they require are part of the rendering process. My framework uses concepts similar to those of the *Painter's algorithm* to create occlusions that are correct for a 3D simulated scene on the perspective ℝℂS. But my algorithm is quite different because it takes advantage of the structure in the perspective tiled floor.

I rejected the traditional Painter's algorithm for my framework because it relies on sorting the 3D polygons relative to the viewpoint. My framework only simulates 3D, and the 3D definitions making the picture are not readily available. The tiled floor provides a natural ordering structure in the perspective ℝℂS, which is at least sufficient to render occlusions to the limit of the tile granularity as this chapter will explain. For comparison, the z-buffer and the Painter's algorithm, are described here. The descriptions reveal that the former is inappropriate and the latter less practical for 2D use than the algorithm presented subsequently.

7.2.1 Depth Buffering

Depth-buffering, often called the z-buffer algorithm, is the most common method used for solving occlusion in the 3D rendering pipeline. Geometric 3D primitive are rendered to pixels on the screen in several processing stages. In the transition from 3D to 2D graphics, primitives generate fragments with colour, depth and texture coordinates. The frame buffer for each pixel of the image keeps track of the smallest depth value: a comparison is made; the color is replaced when the depth value is smaller than the one stored, indicating the stored one is occluded.

Hardware-supported depth buffering is used in most 3D graphics applications, but is not natural to my framework. It works well when the processing order of elements in the 3D scene is uncorrelated with their depth, which goes along with the viewpoint being selected late in the pipeline and changing frequently. But in my case, the scene is conceived as it appears on the canvas: the viewpoint is picked first, the composition developed in accordance. Artists do not assume visibility to be solved by an automatic mechanism provided at the latest stage. A painter considers the scene order throughout, taking definition shortcuts: for the most part[2] he draws only what he sees. Thus, my goal is to determine a sequence in the perspective $\mathbb{RCS}$ that can be used to render scene occlusions.

7.2.2 Painter's Algorithm

The *Painter's algorithm* in computer graphics attempts to handle visibility as painters do. Its name indeed refers to the common abstraction by which an image is conceived back to the front by traditional artists[3]. The Painter's algorithm solves occlusion by sorting surfaces in the scene according to the distance to the eye: the order permits rendering the more distant surfaces before the ones that are closer. Unfortunately, this algorithm, which is log-linear in the number of surfaces to be rendered, works only when a total order on distances exists. In practice, often there is no total order, and refinements on the Painter's algorithm are used [80]. Then the worst running time becomes quadratic.

To see the difficulty consider Figure 7.1, which is rich in occlusions. The difficulty occurs when surfaces overlap all the x, y and z dimensions in the perspective $\mathbb{RCS}$. First, sorting the faces by farthest z (the largest z coordinate among all the face's vertices) produces a front face occlusion order, the brown behind blue behind red, which is both correct and easy to compute. But when there are z overlaps, ambiguities must be resolved: the painted lateral face of the blue block and the brown front face need further tests to determine if occlusion occurs. The tests include pairwise checks on x- and y-extents plus a calculation of the intersection between two planes, all for a determination that is visually self-evident. More difficult cases exist, some of which require polygon splitting, and the result is a conceptually simple algorithm that has $O(n^2)$ worst case running time—where n is the number of polygon of the scene that are not culled—to handle the exceptional cases correctly, belying the simple metaphor on which the algorithm is based.

[2]Structures are sketched transparent, with dashed lines indicating hidden edges.

[3]In fact, the CG Painter's algorithm is an abstraction of artists' practice. Artists optimize by not painting areas that will be occluded by nearer picture elements.

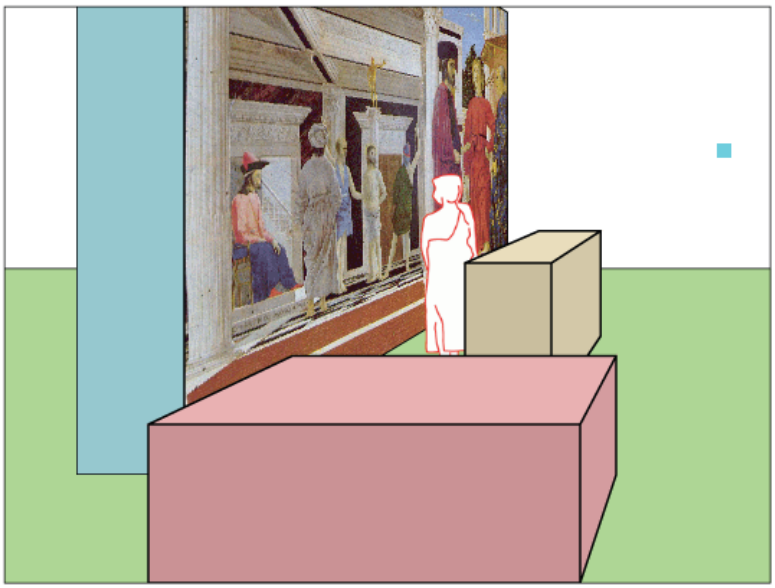


Figure 7.1: Two longitudinal blocks with a figure in between and a block at the forefront. The spatial sequence of the objects in 3D and in 2D is obvious: the blue painted block is behind the panel figure which is behind the brown block, with all of them behind the front pink block. The sequence to draw each face back-to-front determined by the Painter's algorithm on the computer is not as simple; many computations are required.

7.2.3 Discussion: Comparison to artists' method

Artists proceed quite differently from either of the two methods described above. They do not test each point of the scene as does the z-buffer algorithm. And while the Painter's algorithm is metaphorically similar to artists' practice, the two diverge strongly in detail. Artists avoids the many test the Painter's algorithm in practice computes. Artist's visibility computations are in comparison painless though equally effective. When the rare depth intricacies occur they solve them by local considerations guided by experience and geometric intuition.

For example, reconsider the visibility of the scene of Figure 7.1 as conceived by our human senses. We may begin by comparing the scene depth of blocks, but we simultaneously consider the overall scene organization to resolve difficult occlusions. Indeed, scene depth immediately separates the pink block from the other two and we know to draw it last. But different visual judgments determine the occlusion between the blue and brown blocks: the eye sees that to conceive that the blue block is lateral to the brown one, behind it in the perspective $\mathbb{RCS}$, and thus should be drawn first.

This simple example shows that human vision, artistic vision in particular, does not need a total order to depict occlusion. We use many principles in sorting a scene, including general knowledge such as the non-intersection of discrete objects. The ordering approach is *ad-hoc*, solving one problem at a time using a combination of perception, memory and cognition. Occlusion is processed out of sight in early vision, along with other depth cues, providing us with direct apprehension of relative depths of objects along with their other geometric relationships. When drawing, we reconstitute occlusion from geometry, using early vision as an immediate check on correctness. Because the raw material of early vision is, in essence, the perspective $\mathbb{RCS}$, these visual methods transfer directly to drawing on the image plane.

Because occlusion is computed from geometry, many potentially problematic occlusions are limited to juxtapositions of geometrically complex objects, such as among the fingers of a hand or the folds within a cloth. Because such juxtapositions are localized they can be solved locally, which limits complexity. In fact, when analysis fails artists create a model and draw from life.

The artist's top-down approach to occlusion is computationally efficient: faces or polygons greatly outnumber objects; objects are organized to minimize complex occlusion; and where complex occlusion is unavoidable it is usually rendered as a cartoon.

Thus, for the most part it is clear that we have strong and reliable intuitions to effectively order the objects that make up a picture (which is demonstrated below), such practice makes use of interpretations on canvas appearance and the 3D scene relative arrangement of its objects on a qualitative level without precise physical measures, i.e., 3D distances computations are explicitly inexistent. The goal of my rendering algorithm is to take the simplicity of the Painter's algorithm and showing that in the perspective $\mathbb{RCS}$, there exists a simple to implement depth sequence that resolves all depth occlusions, which is closer to human intuitive spatial order.

7.3 The Approach

In my framework, the viewpoint is more or less fixed throughout the picture creation, only limited changes in the eye position are permitted—parallel to the image plane or changing the distance away from it. Neither change alters the spatial total order of the scene elements with respect to the viewer, the relative positions among the scene objects are maintained. In other words, the orientation of the viewer with respect to the scene is heavily constrained; for example the framework does not permit viewing the scene from behind, reversing the depth ordering.

7.3.1 Perspective $\mathbb{RCS}$ Order

Given my context, which is to work directly from the perspective $\mathbb{RCS}$, the z-buffer mechanism is inappropriate. Neither is efficiency the concern, nor viewpoint full freedom. Considering the rigidity of the canvas as a development space for an image, a method that resembles the process followed by the painter who considers sequentiality within the scene the visible groups of the image content is best suited.

My framework uses an adaption of the Painter's algorithm. Instead of ordering each 3D polygon in 3D space, possibly considering splits, the tiled floor structure is taken advantage both globally as well as locally. The intention is to use a more natural decomposition than the one emitted by searching through the entire display list. The tiled floor gives a dominant distinct ordering structure, providing sequential forms that conveniently divides the picture, cutting the scene volume into a small number of discrete rigid sub-spaces, that can easily be ordered. I identify local and explicit splits, rather than arbitrary ones that the Painter's algorithm conceives.

Section 7.5 explicitly studies simple heuristics that reproduce artists ordering of full blocks in context, but those heuristics are challenged as not constituting an automatic deterministic approach applicable in general. Two important questions my rendering addresses:

- Are 3D distances required to calculate visibility? To which extent can perspective organization, its pattern on the canvas, solve the visibility problem directly in 2D?
- Can full objects be sorted? Is it a valid and solvable question? What are the limitations?

The Painter's algorithm does fewer checks than the z-buffer. My goal is to do a minimum number of depth checks, including some types of shortcuts taken by artists, but with the guarantee of a precise sequence for a granularity.

7.3.2 Perspective $\mathbb{RCS}$ Symmetry

Rendering in my framework takes advantage of perspective picture and volume symmetries. The perspective $\mathbb{RCS}$, with a tiled floor defining lines that are axis aligned to the line of sight, creates a picture symmetry important to volume visibility. Thus, to describe the rendering mechanism some geometric terminology within the perspective $\mathbb{RCS}$ setting is helpful.

The **central vertical** of the perspective picture is the vertical line passing through the pvp. While usually centered it can be offset (see Figure D.3). The central vertical and the horizon

divide the perspective ℝℂ𝕊 into four **quadrants** with the *pvp* at the center. The quadrants can be labelled as follows: south-east (SE) and south-west (SW) below the horizon, north-west (NW) and north-east (NE) above it. In terms of the two axes, a projected element of the scene is either fully contained in one quadrant or crosses between quadrants. Placement in and across quadrants determines the appearance of each volume and the order in which its faces are drawn. In particular, opposite quadrants expose symmetrical properties, as the axes defined by the *pvp* are core to the reflective geometry inherent to pictures created in the perspective ℝℂ𝕊.

Since placement influences rendering, giving labels to block faces according to the global placement of the block within the perspective ℝℂ𝕊 helps to explain the rendering sequence. In particular, for a volume it is useful to distinguish its lateral faces, according to their relation with the *pvp*. The **medial** face is the lateral face of a block that points towards the picture central vertical—the one that faces the plane made by the *pvp* and the central vertical as anchored at the baseline. The **dorsal** face points away from it, the dorsal face is directed toward the frame vertical edges.

7.3.3 Overview

One of the benefits of composing on the perspective ℝℂ𝕊 is the existence of drawing sequences that determine occlusion and visibility correctly: the tiled floor gives spatial coherence to the space; in particular because volumes are axis-aligned to its depth grid. Nonetheless, the interest is in determining those drawing sequences that artist follow naturally on the basis of visual intuition. Making artistic practice an algorithmic procedure requires the artists tacit knowledge, their intuition, to turn into deterministic steps that are handled explicitly in the implementation.

The theoretical description of the rendering followed by my framework is done in three stages. First, basic geometry and some simple rules are developed by considering rectangular volumes isolated from one other. Then a few multi-volume examples, where visual intuition naturally and efficiently selects which of several possible rules to apply, demonstrate the difficulty of devising a general algorithm producing the intuitive sequencing results always correctly. Finally, a deterministic solution is presented before showing a representation of a constraint that encodes all the drawing sequences for a scene. The rendering process is based only on measurements in the perspective ℝℂ𝕊 and always creates correct occlusion. The explicit rendering algorithm as implemented in the practicality of my framework is presented in Section D.5.

7.4 Blocks in Isolation

Before considering the global problem of occlusion, it is worthwhile considering it locally. Finding a drawing sequence among the faces of an isolated block teaches many of the concepts that are relevant to solve the general occlusion problem. It also suffices to render a scene of blocks that do not overlap one another. The local visibility of block faces depends on the block's placement, which determines

1. the faces to be rendered, and
2. the sequence in which to render the faces.

The perspective picture and block symmetries described in Section 7.3.2 categorize placements that determine the rendering of a block. The rendering sequence is most simple when only visible faces are considered. Thus, face visibility, the simulation of back-face culling based on patterns on the perspective $\mathbb{RCS}$, establishes a rendering sequence on the set of visible block faces. This basic sequence is the extended to include all the faces of the block.

7.4.1 Face Visibility

Considering only blocks that rest on the floor, the visible faces are the front face, possibly the top face and possibly one of the lateral faces. Figure 7.2 illustrates blocks with different sets of visible faces. The top face is visible only when the block is fully below the horizon. For the left-most block, the right lateral face is visible; for the right-most block the left lateral face is visible. Other lateral faces should be either culled or be overdrawn by other faces: artists draw their construction lines when they are necessary to resolve geometric ambiguities. The middle block has neither of its lateral faces visible. The lateral faces are most likely to be problematic: which are visible, if any? Perception and experience guides us well. The implicit rule we follow depends on the perspective picture symmetry described in Section 7.3.2, specifically the medial and dorsal distinction. All medial faces are visible, while no dorsal faces are.

The two lateral types have further properties important for guiding visibility decisions. Within a volume, the medial face, if one exists, is the closer of the lateral faces to the central vertical of the perspective picture. However, a medial face does not always exist. When a volume spans the central vertical, the two lateral faces are both dorsal, as neither 'faces' the pvp. Thus the two face types determine the visibility of lateral faces.

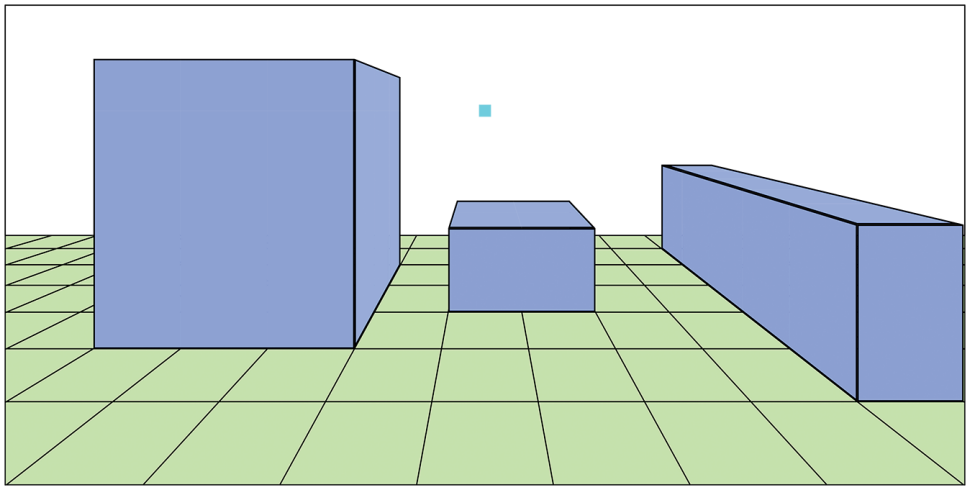


Figure 7.2: Examples of visible block faces. On the perspective picture the visibility of faces depends on the block's placement with respect to the horizon and the central vertical. The front face is always visible. The top face appears only if the full block is below the horizon. Medial faces are also visible.

A similar distinction between the visibility of the bottom and top faces exists. In this case the relationship between the block height, its horizontal faces location essentially, and the horizon is the determinant: the top face is visible when a block is fully below the horizon; the bottom face is visible when a block is fully above the horizon; in both cases, only one of the opposite pair has to be drawn. When a block spans the horizon neither the top or bottom face is visible.

The following set of rules summarizes face visibility for a block in isolation.

- The front face is always drawn.
- Of the lateral faces, only medial faces are drawn. A medial face exists when a block is fully contained in one side of the perspective picture, either east or west. The medial face is the lateral face contained on the orthogonal plane that coincides with the closest orthogonal floor line to the central vertical.
- When a block crosses the central vertical both lateral faces are of dorsal and neither is drawn.
- When a block is fully below the horizon, south, its top face is drawn.
- When a block is fully above the horizon, north, its bottom face is drawn. This case occurs when blocks are stacked up.
- When a block crosses the horizon neither top nor bottom face is drawn.

The three blocks in Figure 7.2 illustrate these rules: all three front faces are visible; the two exterior blocks, each fully contained in one quadrant, have visible medial faces; the top face of the block that crosses the horizon is not visible; the lateral faces of the middle block, which crosses the central vertical, both of which dorsal, are invisible.

These rules are interesting because they reproduce back-face culling without computing 3D normals, using nothing more than the symmetric patterns inherent in the perspective ℝℂ𝕊. Distinguishing between block faces according to their placements within quadrants defined by the *pvp* location, exploits the symmetry relation of the perspective ℝℂ𝕊 and the geometric volumes it contains: central perspective pictures are constructed in a space where symmetries naturally occur.

Other geometric volumes are more complex: dorsal and medial regions of a volume surface can often be roughly defined, but the dorsal side may be partially visible (Section 9.2.1). However, compared to the dorsal part, the medial part is more visible and closer, which is important when establishing a rendering sequence.

7.4.2 Sequence of the Faces

When only the exterior faces, which are visible, are rendered, the order is inconsequential. However, the following back-to-front sequence is natural and generalizes well.

1. Medial.
2. Top or bottom: at most one is drawn, never both, sometimes neither.
3. Front.

To elaborate on the above sequence, it is worth considering inner faces. Doing so provides observations that are fundamental for achieving a rendering sequence that takes into account occlusion among blocks. Indeed, simple geometric volumes have more than three potentially visible faces. Thus, a sequence taking into account all the face partitions—the symmetrical divisions: front/back, top/bottom and dorsal/medial—contribute to the general problem of local visibility within a arbitrary volume.

Figure 7.3 shows occlusions between the vertical faces of a block in the perspective picture, making apparent the overlap between its inner and exterior faces. (This inner/outer exploration contributes to the inter-block sequence as an inside face of a block is co-located with outside face of an adjacent block.) Opened blocks, limited in height, make apparent the inherent

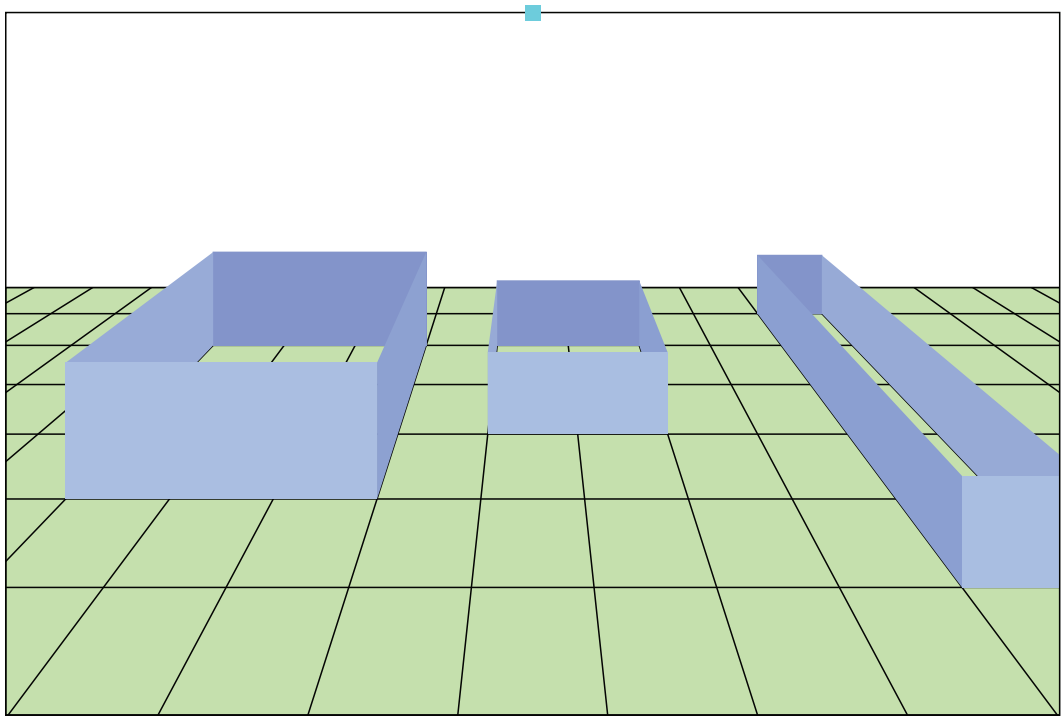


Figure 7.3: Overlaps between vertical inner faces of a block: the front and medial faces overlap the back and dorsal faces. Were the opened blocks taller more face overlaps would occur.

occlusions, those guaranteed to exist. When the top face is omitted, the following occlusions which occur are apparent.

- The front face occludes dorsal face(s).
- The medial face occludes the back face.

And were the bottom face and top face drawn, there would be the following occlusions.

- The top face occludes both back face and dorsal face(s).
- The front and medial faces occlude the bottom face.

Furthermore, if the block were taller, there would be a height at which more occlusions are introduced. Depending on the block's shape, specifically the ratio between its width and its depth, and on its placement in the perspective ℝ℃𝕊 any of the following occlusions could occur.

- The medial face occludes the dorsal one (the rightmost block in Figure 7.3).
- The front face occludes the back one (the left two blocks in Figure 7.3).

Listing these occlusions makes clear a well-defined order that guarantees correct rendering within a block: the front, medial and top faces dominate when blocks rest on the ground plane[4], which is consistent with the visible sequence given on the previous page. The complete sequence is a follows.

1. Back.
2. Bottom if volume is fully below the horizon; top if fully above; otherwise either.
3. Dorsal (possibly two).
4. Medial (if it exists).
5. Top or bottom face not drawn in step 2.
6. Front.

The complete rendering sequence follows a back-to-front order in depth in addition to an order from the perspective picture exterior towards its center lines. The latter renders the lateral face farther from the central vertical first: a dorsal face before a medial one and a horizontal face belonging to a slice farther from the horizon before one nearer. Thus, the following critical directions in the perspective $\mathbb{RCS}$ have been identified:

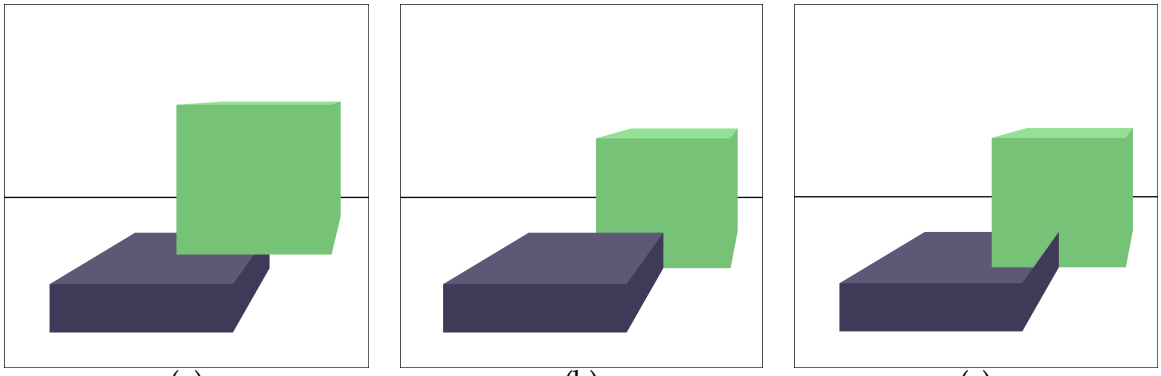
- the directions from the *pvp* to the baseline along the floor orthogonals, for the back and front faces,
- the directions from each vertical frame edge to inwards toward the picture central vertical along the floor horizontals, for the lateral faces, and
- the directions within the $\mathbb{RCS}$ horizontal are from the baseline and the frame top edge towards the either side of the horizon for the top and bottom faces.

These sets of directions suffice for rendering the visibility within faces of a single block. They also support rendering the occlusion among blocks.

[4]The bottom face actively occludes other faces when the block is contained above the horizon, which is omitted in this example.

7.5 Among Block: Artists' Practices

This section examines a variety of artists' solutions to concrete problems, which form the foundation for general rules used for determining visibility in the framework. The algorithm needs to respect the ad-hoc heuristics used by artists to resolve visibility problems locally as they occur.



(a) (b) (c)

Figure 7.4: Drawings (a), (b) and (c) differ only in occlusion yet receive completely different interpretations. Because wrong occlusion is immediately and effortlessly visible, artists learn early what is and is not visible.

Artists, and indeed humans in general, have little difficulty solving perspective visibility problems. They treat each object as an unity and paint them one after another. Disassociating polygons from the object they comprise is exceptional, not the norm. Presented with three drawings in Figure 7.4 the eye immediately sees

(a) a taller block slightly behind and above a shorter block,

(b) a taller block fully behind a shorter block, both resting on the floor, and

(c) an uninterpretable collection of areas.

These interpretations correspond with the gray block being drawn before the green one (a); the opposite sequence (b); and the polygons of the two blocks inter-mixed (c). Artists easily conceive the appropriate sequence according to the context, using quick and effortless visual feedback to learn which occlusions to draw to create the geometry they wish to portray. On the computer the interpretation is not as simple: a completely general solution to the interpretation problem in 3D does not exist, let alone the inverse solution that would allow the ad-hoc visibility calculation that comes so easily to the artist.

The remainder of this section considers two attempts to generalize the simple rendering sequence that correctly rendered the visibility of a block in isolation, to a scene where between block occlusions are present with all blocks resting on the ground plane. It concludes that the problem requires a more complex rendering sequence to achieve correct occlusions deterministically using only 2D image coordinates, which is developed in the subsequent section.

7.5.1 Heuristics

For a block in isolation an occlusion-respecting rendering sequence was established by interleaving the three pairs of opposite faces. Locally the spatial positions of the pairs separate geometrically and are easy to sequence accordingly. From such treatment, it is logical to conceive two possible heuristics that might define a rendering sequence for global visibility.

- Heuristic A (HA): Paint back-to-front and break ties by painting outside in.
- Heuristic B (HB): Paint outside-in and break ties by painting back-to-front.

To apply these two heuristics a point is selected to provide the sequence: the bottom corner point on the front face farthest away from the central vertical, i.e., the picture coordinates of the bottom dorsal corner of the front face. While such choice may seem arbitrary[5] it is the closest to ideal, as the most lateral and front most, it is a location that denotes the separation between the visible and occluded part within a block.

7.5.2 Examples

Simple examples contrast the intuitive approach to visibility provided by human vision and the application of heuristics using the tiled floor priority on its two directions. Artists intuitively perceive the total depth order in a scene in global and local terms and use it to create the image on a flat surface. As the image takes shape, perception naturally and effortlessly sorts which object or part of it that needs to be handled next. Reflexes exercised and learned on simple scenes translate to complex scenes, where a deterministic sequence, with the process relying on heuristics, in context solutions, rather than an explicit algorithm, exists.

Each example (Figures 7.5 to 7.8), shows two collections of blocks on either side of the central vertical: the spatial configuration of each is the reflection of one another about the

[5]Many other fixed points would produce similar results.

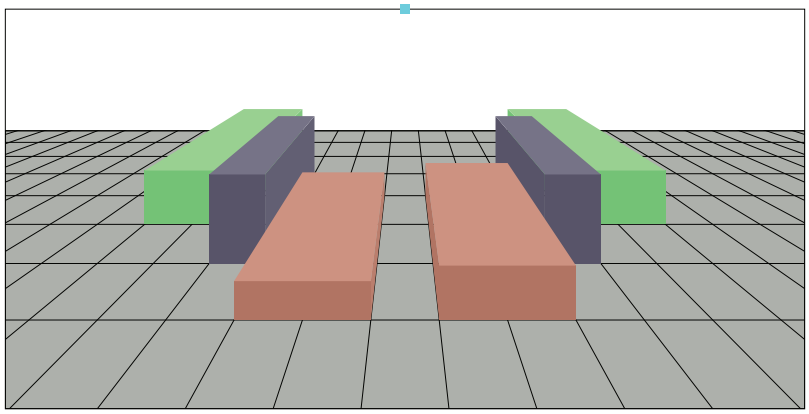


Figure 7.5: On each side of the central vertical a set of three blocks is arranged on the ground plane. Laterally touching, they approach the baseline and the central vertical together, with a clean separation where the central vertical is empty of blocks. Heuristic A is applied to provide the rendering sequence on the west side, while Heuristic B is used for the east: both simulate the correct image occlusions.

central vertical. Heuristic A is applied to obtain the rendering of the blocks of the west side of the image and Heuristic B for the east side. The captions of the figures comments on the success and failure of the heuristics for the particular spatial arrangement of the blocks on the tiled floor.

Figure 7.5 shows that for some block configurations, both heuristics produce the same sequence and the correct occlusions. The two heuristics work because the spatial arrangement of the blocks is most congenial: the deeper blocks are also most lateral and the progression is from dorsal and deepest to medial and nearest.

Figure 7.6 shows a configuration where HA (back-to-front) fails. Each block has a 3D x-extent in addition to a depth range, both of which must be taken into account when considering image plane occlusion. While block painting cannot be sequenced primarily by depth, it is counter-intuitive that lateral positions should always dominate the rendering sequence. Figure 7.7 shows that HB (lateral) also fails, demonstrating the natural insufficiency of the lateral first sequence.

These simple examples suggest that there is no well-defined rendering sequence for full blocks defined by a single heuristic: even simple arrangements may require a mixture of the two heuristics, as Figure 7.8 shows. Each heuristic applies locally to render correct occlusions, but an adequate merge of the two heuristics is needed to get correct global results. Artists

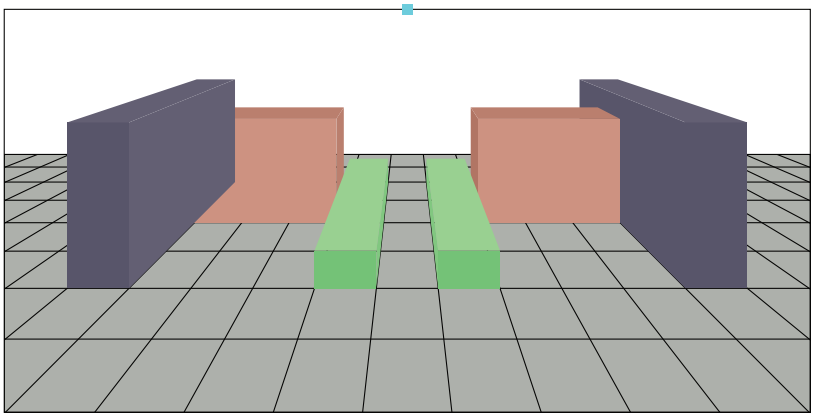


Figure 7.6: On each side of the central vertical a wide block is flanked by two long blocks. On the west following back-to-front first (HA) the image appearance is clearly incorrect: the partial occlusion of the medial face of the dark gray block is missing, the medial face overdrawing the front and top faces of the pink block on its medial side. Artists would automatically identify the dominance of lateral position over depth, painting the configuration as on the east side: HB produces this sequence.

easily conceived such an appropriate and effective sequence.

Those examples illustrate that a universally successful drawing sequence must be more complex than the two simplest heuristics that only work locally and for a particular context.

7.5.3 Lessons

The examples demonstrate three points.

- Neither of the obvious heuristics calculates correct occlusion for all possible configurations; the two conflict when blocks span tiles.
- Each heuristic is sufficient only when artist takes exceptions into account, exceptions that are easy to identify and correct by eye.
- Artistic rendering is efficient: local shortcuts sort objects by intuition.

On the basis of these observations, which contrast the successful drawing sequences of artists with the failure of simple heuristics, I discovered a rendering sequence that guarantees correct occlusion. The tiles of the floor are essential, forming a spatial structure at a sufficiently fine level of detail. Taken literally the heuristics can be interleaved to render the scene correctly.

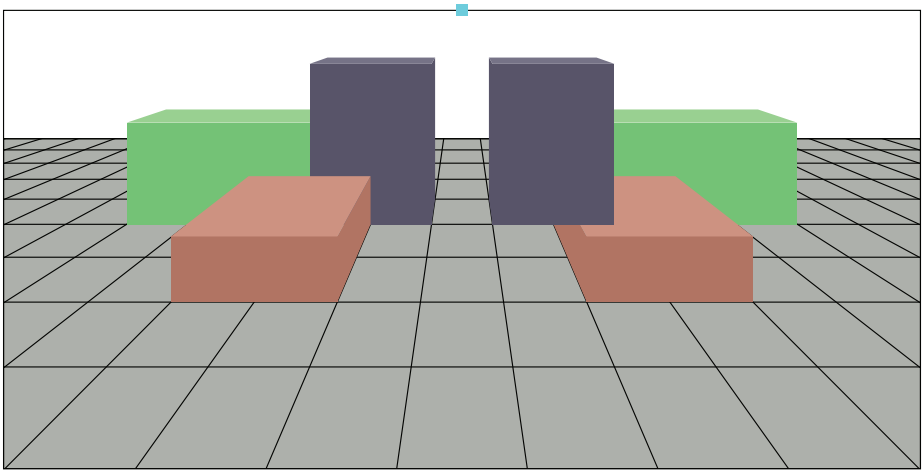


Figure 7.7: On each side of the central vertical two adjacent blocks at the same depth have a block in front of them touching both. On the west side, HA provides the correct result: artists naturally follow the depth first sequence, drawing the two back blocks initially. On the east side, HB, this time, fails as a front face of a deeper block improperly occludes top and medial faces that should have remained visible.

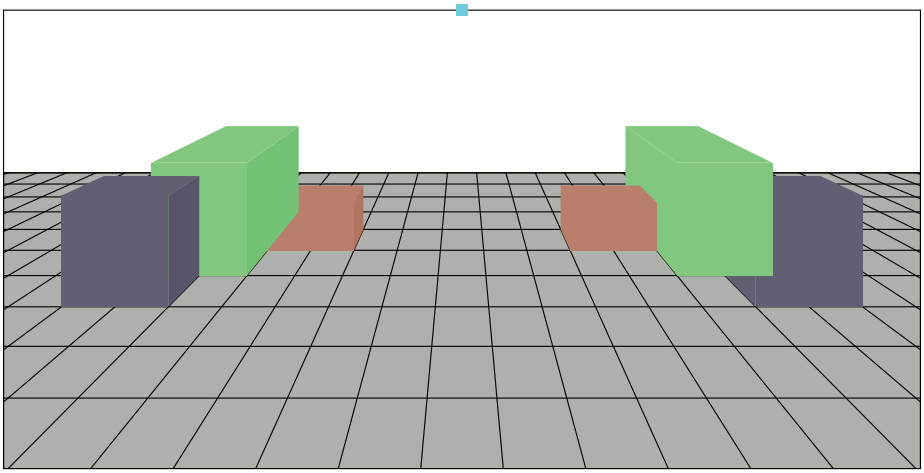


Figure 7.8: Simple block configuration demonstrating the insufficiency of both heuristics. Medial face occlusion is missed by depth first (HA on the west) and front face occlusion is missed by lateral first (HB on the east). Artists easily follow an intuitive sequence: they know to draw the green block before either of the two, the subsequent sequence being immaterial.

7.6 Among Block: A Solution

The perspective ℝCS directions, towards the central vertical from the left and right edges of the picture frame and towards the horizon from the top and bottom edges, sort the distances from scene objects to the eye. They are combined at the granularity of the scene, single tiles, which provide both object definitions and guarantees of correctness. This section presents a correct sequence, one of a large set described in Section 7.7. Which sequence is optimal varies from scene to scene. Here, the specific drawing algorithm used in my framework is explained.

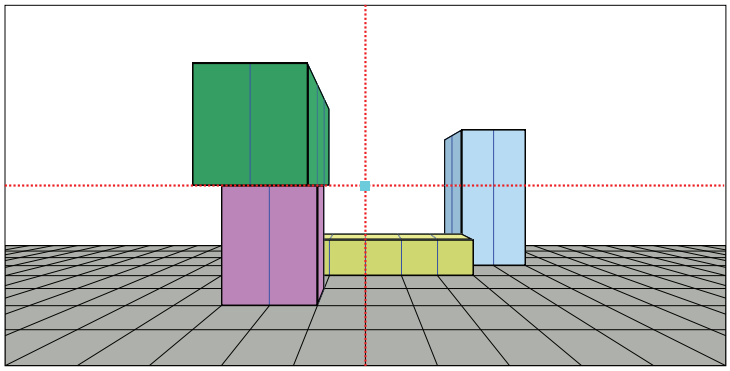


Figure 7.9: Several volumes divided on their unit block boundaries. Each corresponds to a single tile. A volume of any size may be subdivided into a discrete number of unit blocks. The four blocks demonstrate the independence of occlusion between quadrants, an immediate effect of the perspective ℝCS symmetry.

7.6.1 Unit Blocks

Rendering the visibility of scene objects in the perspective ℝCS uses a sequence among the tiles that span the ground plane of the scene. Each tile defines an area in the xz-ground plane, and at the same time a volume extending above it in the y-direction. It is a unique and important feature of the artist's perspective construction that a two-dimensional sequence handles occlusion correctly.

A *unit block* is the part of a block above a single tile: thus the projection of a unit block to the ground plane does not span tiles. A block, as the term has been used, is composed of one or more unit blocks that fill its rectangular volume. Figure 7.9 shows the boundaries of the unit blocks that comprise several blocks. Note that more than one unit block can be above a single tile, as occurs where the green block is stacked above the pink one. Drawing order with

respect to the horizon deals with the unit blocks contributed by different blocks above the same tile. Among the collection of the unit blocks that comprise a scene, a rendering sequence that guarantees correct occlusions follows an order inherent in the floor tiling.

7.6.2 Quadrant Independence

It greatly simplifies the problem that the four quadrants can be rendered independently, making it possible to establish a separate sequence for each quadrant. Specifically,

- blocks left of the central vertical cannot occlude blocks to the right of the central vertical and *vice-versa*, and
- blocks above the horizon cannot occlude blocks below the horizon and vice-versa.

The *pvp* location in the central perspective schematic representation specifies a separation of planes, which can be taken advantage of. The impossibility of occlusion is illustrated in Figure 7.9. The pink block, fully contained in the south-west quadrant, can occlude the part of the yellow block that is in its quadrant, but cannot occlude the green (NW) or blue (east side) blocks. The yellow block spanning SE and SW can occlude the pink or blue blocks, but no block that is above the horizon. The green block can only occlude blocks in NW, here the empty set. Thus, the sequence in which a scene is drawn has quadrant independence: a sequence in each quadrant has to be followed but interleaving arbitrarily between them is possible. This freedom allows the programmer to create a sequence that optimizes, for example, comprehension or efficiency.

Thus, without loss of generality, the detail to establish the sequence within a single quadrant is given. The following description starts by sequencing the south-west quadrant domain and then gradually integrate the interleaving from the other quadrants in a global sequence.

7.6.3 A SW Quadrant Sequence

Prioritizing among the unit blocks on the ground plane of the south-west quadrant is accomplished by combining the direction from the picture edge towards the central vertical with the direction from the *pvp* to the baseline to produce the sequence shown in Figure 7.10. Tile 0 is the most lateral and deep, while tile 1 is less lateral but at the same depth and tile 5 is similar to tile 0 laterally but less deep. Tile 14 is the closest to the viewer. A unit block at tile 0 does not occlude a unit block situated at any of the other tiles. Similarly, no unit block at tiles 0–13 can

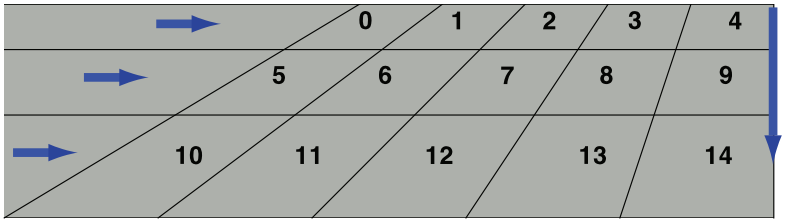


Figure 7.10: An order by which to process the tiles of the south-west quadrant to guarantee correct block occlusion. The numbers attribute a back-to-front sequence to the tiles.

occlude a unit block at 14. The tiles by their geometrical definitions correctly intermix the 3D extents of each block in the given perspective geometry. Taking this into account a sequence that draws the scene with proper occlusions is easy to find.

Let us reconsider one of the previous example, the south-west quadrant of which is shown in Figure 7.11. The example contains three blocks, two of them spanning four tiles in depth while the central block spans three tiles laterally, has eleven unit blocks to order. Traversing the tiles of the floor according to the sequence of Figure 7.10, each unit block is rendered in order, which results in the rendering sequence in Figure 7.11. This process draws part of the gray block before the brown one, then completes drawing it later. In this case, of course, splitting is unnecessary: the drawing order grey-brown-green produces correct occlusions. Block splitting is necessary, however, because it produces a simple, robust painting order that produces correct occlusions without requiring geometrical calculation specific to a particular scene.

Figure 7.12 shows the sequence used to draw an arrangement of blocks that are stacked up. This configuration includes several unit blocks over the same tile as well as gaps between unit blocks and the ground plane. Each tile maintains a list of the references to each of the unit blocks above it, in order of their distance from the ground plane. For a structure that is completely contained in a southern quadrant, as is the case in Figure 7.12, the unit blocks above each tile are drawn from the ground plane to the horizon, so as to have the top face of the highest block drawn last.

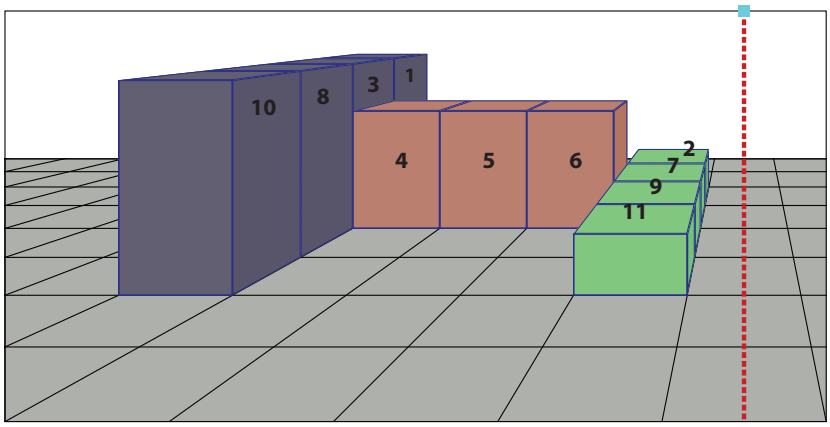


Figure 7.11: Previous example, Figure 7.6, revisited using unit blocks. Thanks to the granularity provided by unit blocks, the conflicts among blocks that span tiles disappear. A spatial order based on the tiles provides a sequence which produces correct occlusions.

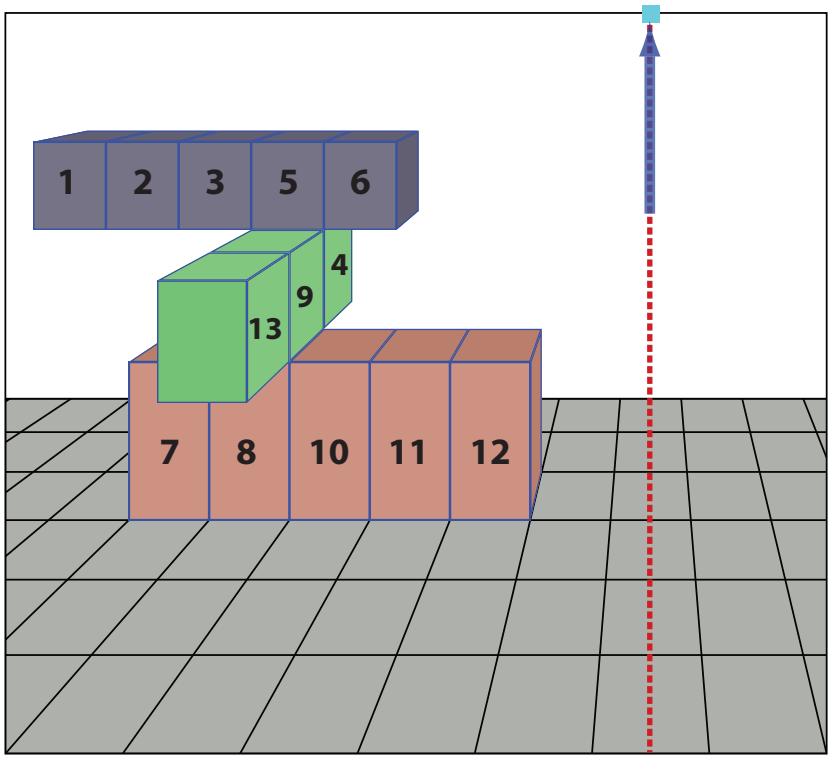


Figure 7.12: Several unit blocks above a same tile. For a pile under the horizon, the correct sequence draws unit blocks above a tile from the ground plane to the horizon.

7.6.4 A Full Sequence

A sequence drawing correct occlusions within the SW quadrant has been given. A similar sequence works for the SE quadrant. However, while the division between the east quadrant and west one can be clean—no unit blocks cross the central vertical—it is not always the case: floors that have an odd number of horizontal tiles at the baseline include a column of unit blocks crossing the central vertical. Figures 7.12 and 7.11 avoided unit blocks in both east and west quadrants. A brute force solution, where each quadrant is drawn independently and the quadrants then merged, requires clipping at quadrant boundaries to achieve correct occlusion assuming back-facing faces are drawn. Alternatively, it is more efficient and elegant to interleave the sequences from the start.

In effect, the tiles apart from the central vertical can be directly intermixed to provide the drawing sequence of the full scene. Figure 7.13 shows an order in which to consider each tile across the east and west side of the floor. Such an approach draws unit blocks that cross the central vertical only once.

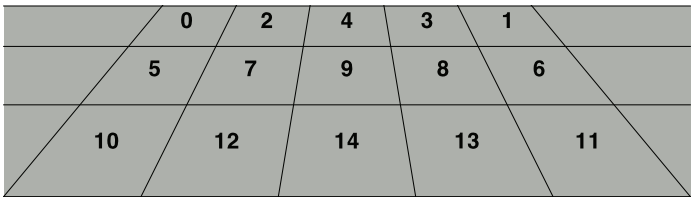
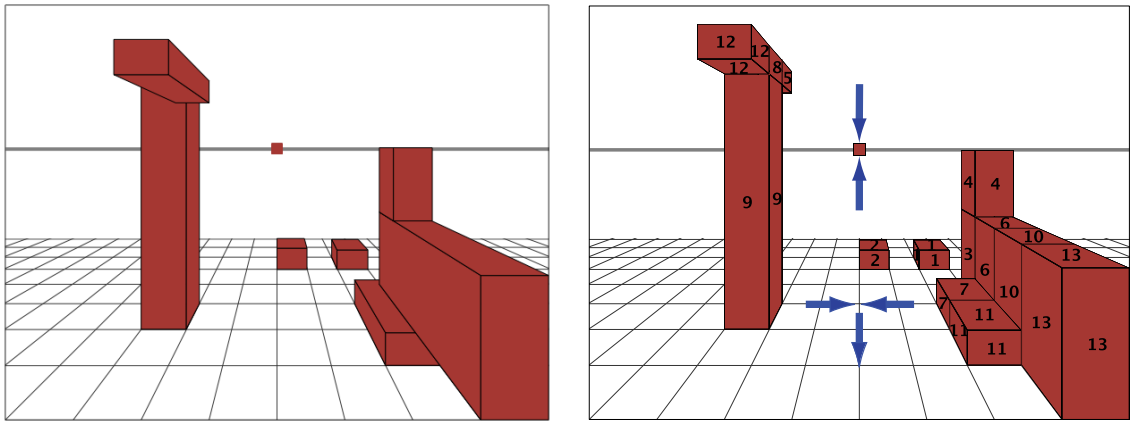


Figure 7.13: A tiled floor with an odd number of tiles along the baseline. To avoid clipping to the quadrant when drawing unit blocks that are along the central vertical, the east and west quadrant tiles are interleaved in a global sequence.

The only remaining problem is including unit blocks in the northern quadrants: unit blocks above the horizon or crossing it. Inverting the reasoning used for the stacked up blocks in Figure 7.12, the horizontal face that matters for blocks above the horizon is the bottom face, as described in Section 7.4.2. In fact, for several unit blocks above the same tile, the correct sequence is from the ground plane to the horizon, interleaved arbitrarily with a sequence from the upper frame edge to the horizon with any unit block spanning the horizon drawn last.



(a) (b)

Figure 7.14: An arrangement with stacked blocks. Simple heuristics do not handle properly the occlusions, as shown in (a). The method that orders using unit blocks renders with correct occlusion, as shown in (b).

Thus, a global sequence on the tiled floor is now known: adding to the ground plane directions an ordering in the space above it. The directions from the horizontal frame edges to the horizon on either side, distinguish distances to the horizon among stacked unit blocks. Figure 7.14 shows a scene with a complex block configuration, contrasting (a) the depth first heuristic of Section 7.5 with (b) the sequence described here. In Figure 7.14 (b) the staked blocks, due to the unit block drawing sequence, have either their top face or bottom one visible according to their positions relative to the horizon. Blocks that span the horizon, the unit blocks of which are drawn last within the stack on a tile, automatically get correct occlusion.

Quadrant independence increases flexibility. Furthermore, separating the drawing sequences eases comprehension and emphasizes the image plane pattern inherent in the geometry of the perspective $\mathbb{RCS}$. But most important independence serves to mix the contributions of each in a global sequence that is optimal for a given scene. The presentation above is an example. The precise order chosen is a matter of convenience.

To exploit the variety of possible sequences it is essential to enumerate the set of sequences within a single quadrant. The next section studies the drawing constraints that exist among unit blocks, defining the set of all correct sequences. The more abstract reasoning justifies the practical solution just given. It shows that the tiled floor has a structure that contains a well-defined set of sequences. This property extends the spatial properties of a perspective scene, which has significance for many rendering problems in computer graphics.

7.7 Among Block: All the Solutions

The perspective ℝCS has geometric properties inherent in the structure of its tiled floor: many drawing sequences exist, each of which renders the scene occlusion correctly. These sequences are derived by recursively applying a single drawing constraint that cascades across the tiles of the floor in a given quadrant. This section describes the drawing constraint that orders the unit blocks of a quadrant in a representation that provides all valid drawing sequences, only considering blocks that are on the tiled floor (omitting stacks as the sequence of blocks within a tile just nests inside the tile order). Among the possible solutions exist sequences that are more effective than the systematic one described in Section 7.6. The global representation of the constraint makes it possible to identify the sequences used by artists, including the shortcuts described in Section 7.5. Describing the set of correct sequences, both completes the theory underlying my framework rendering and connects to other domains where classifying and sorting are also the central task.

7.7.1 A Constraint Problem

The set of orders allows intermixing of the two floor directions in a quadrant. The worst case—the entire quadrant being filled with unit blocks, one on each tile—determines all possible sequences for any scene configuration in the quadrant. It is necessary to formalize for a quadrant arbitrary tile, the set of unit blocks that must be drawn before the one on that tile. In Figure 7.15 the yellow unit block immediately occludes a unit block behind it in the same column and one beside it on its dorsal side. Those relations, applied recursively make the green area of Figure 7.15 the region whose elements should be drawn prior to the yellow block. By a similar argument, unit blocks in the red area must be drawn after the yellow unit block.

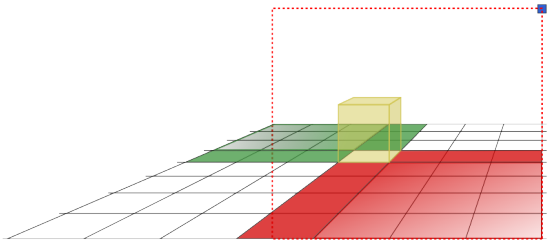


Figure 7.15: A unit block, arbitrarily placed on the tiled floor south-west quadrant. The red tiles contains unit blocks that must be drawn after the yellow block; the green contains those that must be drawn before.

Formally, for unit blocks on the tiled floor, where each is referred by its tile location (n, m), the mechanism is as follow resulting in two main rules. An unit block at (n, m) necessarily occludes unit blocks at $(n-1, m)$ and $(n, m-1)$. Thus unit block at (n, m) must be drawn after unit blocks at $(n-1, m)$ and $(n, m-1)$. *'Drawn after'* is transitive: if unit block at (b, b') must be drawn after unit block at (a, a') and unit block at (c, c') must be drawn after unit block at (b, b'), then unit block at (c, c') must be drawn after unit block at (a, a'). Therefore, the first rule is that all unit blocks at tiles $(i \leq n, j \leq m)$ must be drawn before unit block at (n, m).

Substitute *'occludes'* for *'is occluded by'*, -1 for $+1$ and *'drawn after'* for *'drawn before'*, and a similar set of necessary truths produce the second rule: all unit blocks at tiles $(i \geq m, j \geq n)$ must be drawn after unit block at (n, m).

It remains to decide the order in which to draw unit blocks at $(i > n, j < m)$ and $(i < n, j > m)$, the tiles that are uncoloured in Figure 7.15. The 'problem' is to enumerate all drawing sequences that obey the *'draw after'* constraint. Doing so shows that, given a configuration of blocks, some tile drawing sequences are better—closest to artistic practice where intuition leads to an effective spatial order—than other sequences.

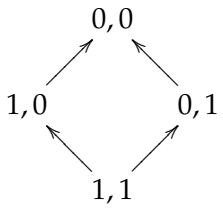
7.7.2 Initial Drawing Constraints

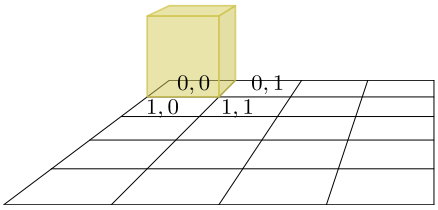
Consider as an example the first unit block at $(0, 0)$ to be drawn, the most dorsal one in the deepest row. It has nothing behind it or on its dorsal side. For the remainder of a SW quadrant with N tiles in each row and M in each column, there remain $NM - 1$ tiles to order.

Occlusion is formalized using the symbol $\prec$, representing the transitive relation *'drawn before'*.[6] For a unit block at $(0, 0)$, we know $(0, 0) \prec (0, 1)$ and $(0, 0) \prec (1, 0)$. Face juxtaposition makes it possible to deduce the drawn before relation from perspective $\mathbb{RCS}$ appearances: a block medial face at $(0, 0)$ is collocated with the dorsal face at $(0, 1)$, which is not visible. In Section 7.4, dealing with the block in isolation, we saw that the occluded dorsal face must be drawn before the medial one, thus by extension of face collocation $(0, 0)$ must be drawn before $(0, 1)$. In the case of a unit block at $(0, 0)$ being higher than the one at $(0, 1)$, part of the former medial face remains visible, which is handled correctly because drawing $(0, 1)$ after $(0, 0)$ occludes exactly the correct fraction of the medial face at $(0, 0)$, leaving the previously drawn upper part of the face visible. Similarly, the front face at $(0, 0)$ and the back face at $(1, 0)$ coincide so that $(0, 0) \prec (1, 0)$. Continuing, $(1, 1)$ may have occluding faces coinciding with faces at $(0, 1)$ and $(1, 0)$. Thus, $(0, 1) \prec (1, 1)$ and $(1, 0) \prec (1, 1)$ and by transitivity $(0, 0) \prec (1, 1)$.

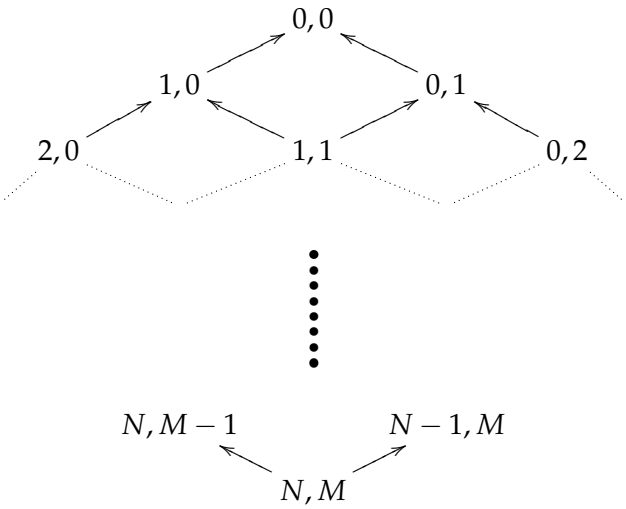
[6] $A \prec B$ means 'draw anything you find on tile A before whatever you find on tile B'.

The drawing relations between those four tiles (as shown below) can be represented by the tile graph, a directed graph, shown beside it:





The occlusion relation between adjacent tiles progresses throughout the quadrant to the front-most tile, (N, M) whose unit block if it exists nothing occludes. Thus, the above tile graph extends to a tile graph that has a single leaf at the lowest level.



7.7.3 Arbitrary Occlusion Constraints

To fill the ellipsis in the tile graph above, reconsider the drawing order for a unit block placed on an arbitrary tile (n, m). Figure 7.16 shows a magnified version of Figure 7.15, with direct drawing constraints at (n, m) indicated by arrows: the neighbouring tiles may have blocks whose faces are collocated. Formally, these drawing constraints are the following:

Farther	front/dorsal	$(n, m-1)$	$\prec$	(n, m)	$\prec$	$(n+1, m)$	(A)
Nearer	medial/back	$(n-1, m)$	$\prec$	(n, m)	$\prec$	$(n, m+1)$	(B)

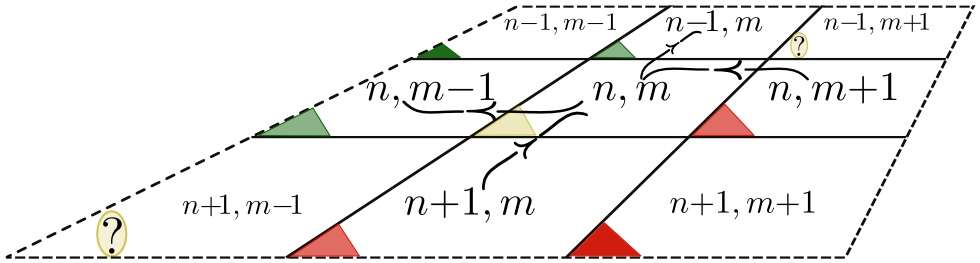


Figure 7.16: Direct drawing constraints at an arbitrary tile (n, m). Neighbouring tiles are labelled with respect to their relation to (n, m). Colour indicators at the corners identify their group: drawn after with red, drawn before with green and indirect constraints yet to be determined with a question mark.

As those direct drawing constraints, takes L-shape paths throughout the tiled floor, an indirect drawing constraint can be generalized as follow

$$(n-1, m-1) \prec (n,\ m) \prec (n+1, m+1).$$

As shown in Figure 7.17, the tiled floor can be taken as a set of diagonals, given by (n, m) with $n - m$ constant, with indirect drawing order constraints along the diagonal (the diagonal points to the dp on the same quadrant). This set is crossed by the other set of diagonals, given by (n, m) with $n + m$ constant, along which there are no drawing constraints (tiles along a same diagonal towards the dp on the opposite quadrant have no occlusion among them).

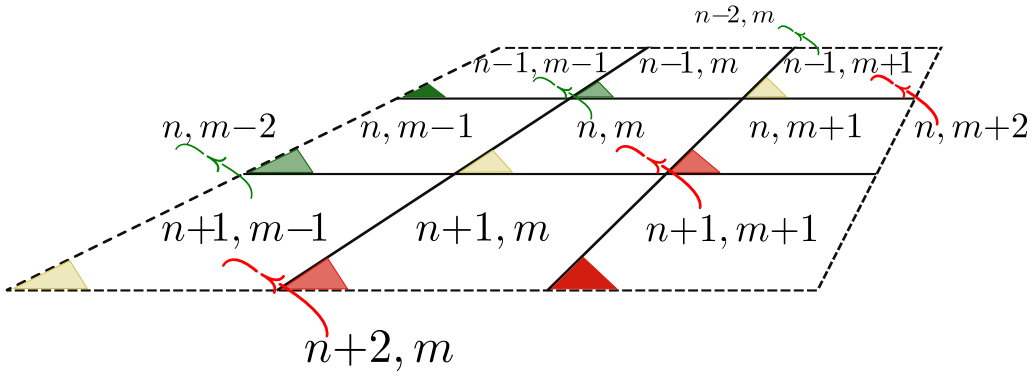
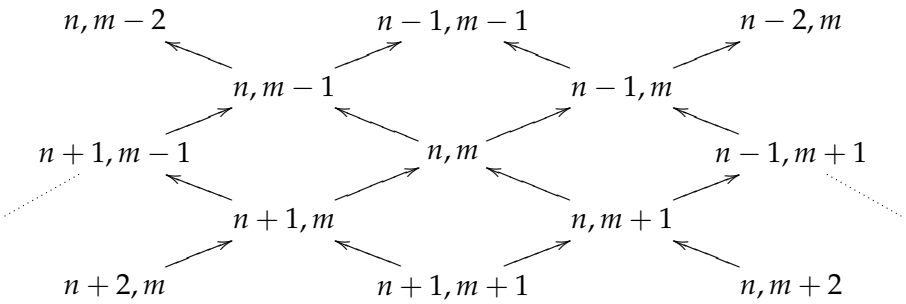


Figure 7.17: Indirect constraints surrounding the (n, m) tile: they follow tile diagonals towards the exterior of the perspective picture.

Rotating Figure 7.17 to make the diagonals horizontal and vertical gives the following tile graph of the constraints around (n, m).



The tile graph, which is easily extended to the edge of the floor, has the following features.

- Nodes connected by a right arrow are on the same row of the tiled floor.
- Nodes connected by a left arrow are on the same column of the tiled floor.
- Nodes lying on horizontal rows have no occlusion relations.
- Nodes in the same column of the tile graph are tiles with corners on one diagonal, with a drawing order constraint among them.

Beyond analyzing potential occlusions I have discovered independent subgroups within which no occlusion occurs (Figure 7.18), which could provide parallelism during rendering.

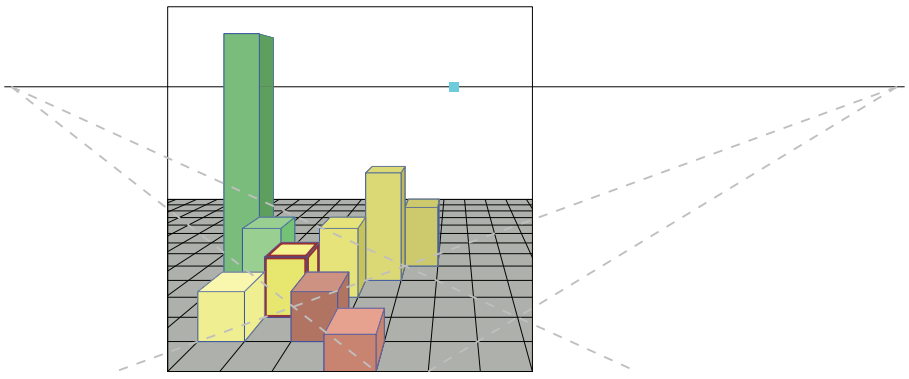


Figure 7.18: Unit blocks lined up on the two types of diagonals. Within the yellow group, there is no occlusion. On the other diagonal, the yellow block in the middle partially occludes the green one directly behind it, both green are drawn before; while the red one directly in front of it and drawn after occludes the yellow block.

7.7.4 Overall sequences

Knowing the constraints around tile (n, m), all drawing sequences for the tiles within a quadrant can be created. For the small example shown in Figure 7.19, the mapping of main traversals of the tile graph onto the physical tiled floor is demonstrated. Figure 7.19 juxtaposes the physical representation of an area of 3×4 tiles with the drawing constraints in the tile graph for the physical floor in the south-west quadrant of a perspective picture.

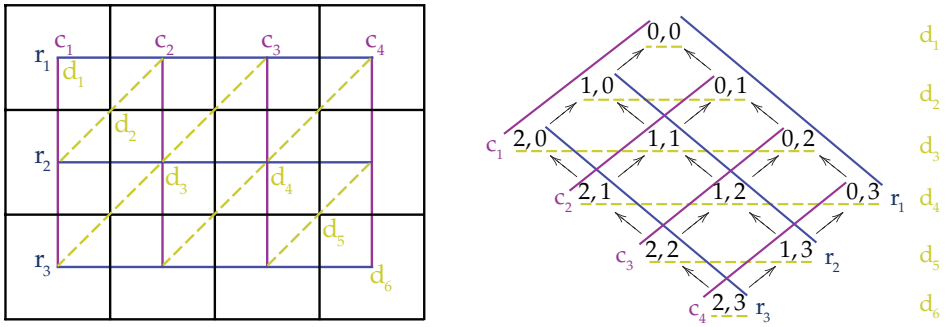


Figure 7.19: (a) Physical tiled floor within the south-west quadrant. (b) The tile graph for the drawing constraints among unit blocks on those tiles: three traversals along floor directions, which give correct occlusions, are highlighted.

Figure 7.19 shows that there are a least three simple traversals that produce correct occlusions for any configurations of blocks on the tiled floor. The tile floor can be traversed in any of the three following ways (each a distinct colour).

1. Level by level, as indicated by the yellow dashed lines: from d_1 to d_6.
2. Diagonally, following the left arrows, to which the purple lines are parallel: from c_1 to c_4.
3. Diagonally, following the right arrows, to which the blue lines are parallel: from r_1 to r_3.

My framework implementation uses the third order, processing unit blocks on tiles row by row from deepest to closest, from the most dorsal tile to most medial one on each row.

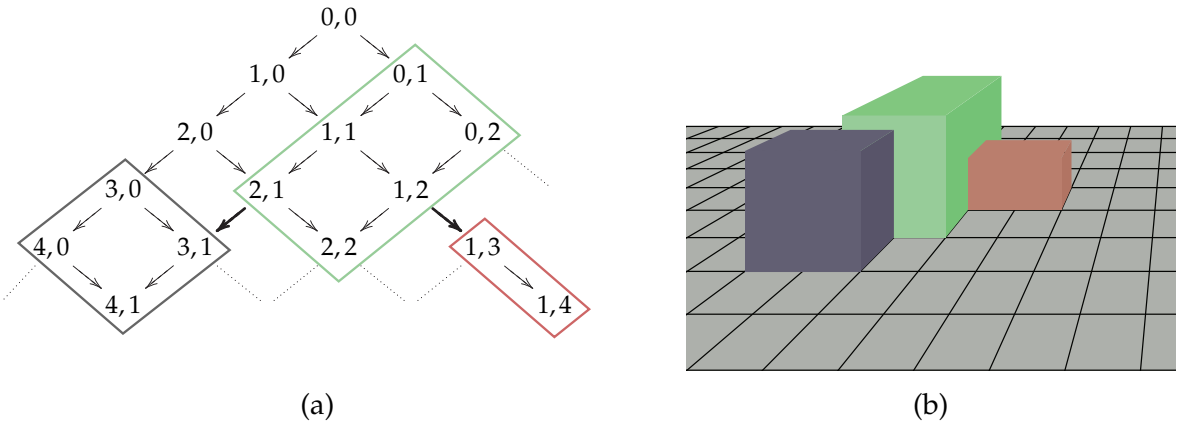
7.7.5 Reduction

At this point, it is worth noting that this problem touches on a variety of deep mathematical problems. The essence is the following. Knowing nothing of the scene that is to be rendered

we are indifferent to which of the $O((MN)!)$ paths are chosen. Any one of the many possible drawing sequences produces correct occlusion, regardless of the occupation of tiles. The every possible drawing order is an example of a rectangular Young tableau[7].

Perhaps representation theory can help us make sense of the following observation. Consider the example of Figure 7.20 (b): an artist faced to this problem would notice that there is an order in which the blocks can be painted—green first, then grey and brown in any order. Can this order be discovered automatically? Figure 7.20 (a) shows that the order perceived intuitively by the artist appears as the tile graph decomposing into subgraphs that can be ordered instead of ordering individual tiles.

The strength of the many drawing orders is that each is correct regardless of the scene. Taking advantage of the strength requires giving it away by letting the scene choose the best drawing order for it. Thus, in Figure 7.20 case following painting in any order the tile content of $[(0,1)\cdots(2,2)]$ by painting $[(3,0)\cdots(4,1)]$ followed by $[(4,3)\cdots(1,4)]$. Discovering subgraphs without human intercession, and finding rendering orders that take advantage of them is an avenue for future research.



(a) (b)

Figure 7.20: A previous example, Figure 7.7, revisited. The unit blocks making up each block are considered as independent subgraphs, connected back together: the two constraints that remain, bold arrows, show that the green block must be drawn first. The order chosen intuitively by an artist can be identified within the directed graph representation.

[7]Young tableaux were originally used by mathematicians for classifying group representations [43]. They have subsequently been studied in theoretical computer science with regards to sorting algorithms [67]. This connection is not surprising, given that occlusion algorithms often rely on sorting.

7.8 Discussion

The rendering algorithm showed that the *pvp* in a central perspective image picture divides the scene ordering into four independent quadrants: the placement of an object on a floor tile matters because of the position of the tile with respect to the *pvp*. The explicit algorithm implemented for my framework (§ D.5) produces predictable occlusions between panels and their surrounding blocks: the order is simple with the picture simulating a schematic representation of a 3D scene; the user is responsible for adjusting the location and a attachment points of panels to produce the desired occlusion.

Drawing the unit block on a tile precedes drawing the panel(s) on that tile. This order ensured that panel content occludes the unit blocks on its dorsal side while it is occluded by those on its medial side. An object is not actually supposed to be inside another one or entangled together. But since the panel attachment is central and the object represented possibly using a semi-transparent panel, it is not uncommon that it is within a tile that a block occupies. However, this particularity of drawing the panel subsequently to the volume permits connecting local occlusions. (In Section 8.2.2 the reproduction of *Saint Jerome in his Study* uses the flexibility provided by manipulating the attachment point of panels that are surrounded by volumes.)

In practice panels are sized to interact visually only with adjacent elements. A user learns to take advantage of the simple and intuitive occlusion rules of the algorithm. The unique control that anchors the panel to the tiled floor makes it possible for the user to place panels within volume gaps and free space: above, below, behind or in front but also in between. The granularity of the tiled floor can be adjusted to provide the level of detail desired by the user.

Equality in depth between overlapping panels violates the assumptions algorithm, as it implies that their real volumes would inter-penetrate (§ D.5.1). Artists, however, who wish to create paradoxical volumes are not prevented from doing so. Judicious use of transparency, adjustment of attachment point and panel rotation can introduce a variety of effects, which future work should make systematic, all the while maintaining the primacy of the 2D picture over fidelity of 3D optical physics.

7.8.1 Correspondence with Artistic Practice

The similarity of my approach to artistic practice should be emphasized. The spatial complexity of the scene is moderated by the symmetry of the central perspective picture configuration.

Artists consider objects locally, in relation to the content immediately surrounding them, treating them in separation of the rest as much as it is possible.

Artists do not work globally at a small granularity. Rarely do they consider order among pixels or polygons independent of the objects that comprise them. They encourage handling image parts at the object level whenever possible. Thus, compared to artistic practice, my algorithm does excessive over painting, as does the usual Painter's algorithm. But, owing to the perspective geometry of my schematic representation the space is pre-sorted, thus my rendering algorithm does not require to globally sort as the Painter's algorithm does. The visibility among picture elements is handled progressively following grid cells.

The object boundaries are fundamental, as are the placements of the main plane. For example, artists paint object interior locally, with only approximate reference to the geometry of the overall scene. The self-occlusions of an object are portrayed precisely with respect to the object frame, mostly disregarding the global rendering setting. (§ 9.2.3 considers an extension to provide such refinement.)

7.8.2 Entanglement Limitations

Inter-penetrations and entanglements between objects are only rarely encountered in nature. They occur for example when hands join with finger-interlocking, in a tangled garden, when the wind blows on light fabric. Artists tend to treat such instances of mingling as exceptions, handled as special cases.

Modelling two hands individually in 3D, then interlocking the fingers naturally is hard for a computer. Denying the artist the opportunity to draw from life in favour of automated modelling and rendering can be perverse. It puts on the computer a task it does poorly, ignoring available user skills.

The distinction between image overlaps that separate easily by objects and those that do not are obvious to the eye. For an algorithm to retrieve the natural and planar juxtapositions involved and needed is less simple. For example, a person appearing through an half-open door is a situation that separates in three panels: the door frame, the actual door and the figure appearing through. A slightly different example, an arm wrapped around the shoulder of another figure, should not be created as two distinct panels, but requires the treatment of the pair within an unique panel. My framework and its rendering trusts the users, who do the simple decomposition but the framework then assists to compose the parts together in the perspective picture. The attachment point permits the former type of interaction: a bird appearing through a window in Figure D.7; the next chapter shows other examples.

Chapter 8

Results & Discussion

The practical goal of my thesis is to assist the geometrical composition of realistic pictures. The framework on the picture surface provides a coherent perspective for architectural elements. It also provides guidance for the placement of the foreground elements having local perspective. The user manipulates the architectural and foreground elements in 2D. Projected shadows provide additional depth cues for spatial understanding.

This chapter presents another flavour of the results of my research: its application to create compositions. I developed construction line algorithms in 2D, projective algebra mappings for working in 2D, and a rendering algorithm with correct occlusion in 2D. The primary result of each is its contribution to the correct functioning of my framework, as this chapter demonstrates. Each has other implications, which were discussed in earlier chapters.

In this chapter, I demonstrate that my framework achieves my thesis goal by mimicking the compositions of several Renaissance paintings, by analyzing the created compositions compared to the originals, by presenting shadow examples situating the picture elements and by exploring a more abstract handling of space following a modern painting. Those schematic representations show that making pictures with perspective can be a 2D task for the user.

8.1 Overview

This chapter presents schematic representations created with my framework. First, a typical result exposes the potential of my framework to combine simulated 3D with 2D composition. Then, a sequence of schematic representations focuses in turn on a variety of capabilities of my framework. A comparative example, two picture geometries for the same narrative, is

used to identify attributes affecting the quality of a composition. Lastly schematic images with projected shadows that contribute to situating objects in depth are presented, followed by a schematic representation that demonstrates a playful use of perspective, created by manipulating construction lines. All the pictures I made are intended to reproduce a feeling similar to the compositions that inspired them: they test my framework and need not to be exact reproductions; if desired they could be improved upon with my current framework.

The schematic representations created with my framework expose its capabilities. Most of the geometries contained in my schematic representation have origins in Renaissance paintings, some of which were discussed in Chapter 3. In the past, art historians have analyzed Renaissance constructions, including paintings in this chapter [65, 102], using paper and pencil. My framework assists such exploration tasks. The schematic representations demonstrate that my framework provides perspective in two dimensions: the user manipulates 2D points, the framework performs 2D evaluations. Approximations to foreground objects and to synthesized architecture reveal the realistic spatial organization of the pictures.

In many cases (Figures 8.5 (b), 8.8, 8.11 and 8.14) I deliberately use generic foreground figures—chosen from the off-line content in Figure 8.1—because the geometric composition depends on the existence of a figure at a position, and not its exact shape. Therefore, generic figures separate the composition from the narrative.

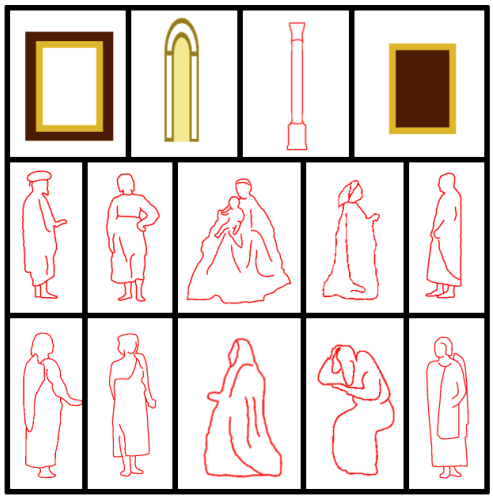


Figure 8.1: Examples of off-line content used in panels. Some I created as delineations of figures, filled in white with transparent pixels outside. (The delineations were traced from foreground figures of Renaissance paintings using Inkscape and converted into semi-transparent raster images using Photoshop™.) Others are simple textures created to decorate architecture.

8.2 An Illustrative Result: *The Trinity*

Figure 8.2 (a) is a schematic representation created with my framework, based on *The Holy Trinity* of Masaccio. Talbot investigated this painting to demonstrate that its composition and perspective adhere to his pattern [102]. My picture shows placement and occlusion correctly with my framework's 2D calculations.

The narrative present in Figure 8.2 (a) is Christ, in death, united with God the Father, observed by four people (Mary and Saint John, plus two Florentines). The scene occurs in an atrium, which is modelled with walls constructed according to § D.2.3. Within panels are the Masaccio figures in silhouette forms: the Florentines and the saints are placed in profile on panels, while Christ and his father are represented frontally. The surfaces on which the figures stand are constructed using blocks. It demonstrates how my framework assists in constructing an explicit perspective: volumes and panels were arranged in the picture space.

(a) (b)

Figure 8.2: *The Holy Trinity* (c. 1426) by Masaccio beside the schematic representation I created using my framework.

Using my framework I created Figure 8.2 (a) by the following steps.

1. I selected a combination of Talbot's planar pattern: one wide by two high.
2. I placed Masaccio's painting, uniformly scaled, as a backdrop to the pattern, either visible as a guide or turned off.
3. I selected a perspective floor appropriate for the foreshortening of the painting architecture and sufficiently discrete for modelling. Figure 8.3 (a) illustrates to this step.
4. I used volumes to approximate the architecture: the forefront volumes on which the figures stand were easy and done first to reserve space behind (as illustrated in the close up image of Figure 8.3 (b)); then the nave was created with pillars supporting the arches (as illustrated in the second image of Figure 8.3 (b)).
5. I roughly arranged the panels to match the cartoons of the painting (Figure 8.3 (c)).
6. I adjusted the panel attachment points to obtain correct occlusion and I moved and sized the panels in accord to 3D and 2D sensibilities (Figure 8.3 (d)).

Both the backdrop of Masaccio's painting and Talbot's pattern were used as models for 2D composition. Creating this representation with the planar pattern guidance reveals its composition organized with an isosceles triangle: on its base the two praying figures, Mary and Saint John on the ascending sides, Christ's head at the apex. This up and inward triangle is reminiscent of the downward one organizing *The Baptism* composition, symbolizing the Trinity (Figure 3.6 (a)). In *The Holy Trinity*, Christ on his cross also suggests this triangle.

Masaccio's painting is not rigidly unified by a single perspective: compromises exist in the perspective of the nave likely owing to the demands of 2D composition [37, 65, Chap. 3, p. 17–21]. My framework picture is governed by an exemplary central perspective: the *pvp* is on the mid-line at the level of the feet of the two Florentines. The picture's geometric center is at the knees of the Christ.

This typical example, which merges dominant figures on panels with architectural elements arranged in the scene depth, supports my thesis statement: my framework enables creating a 2D composition within simulated 3D. This chapter's remainder presents schematic compositions that go deeper into my framework, addressing in turn specific potentials of the framework.

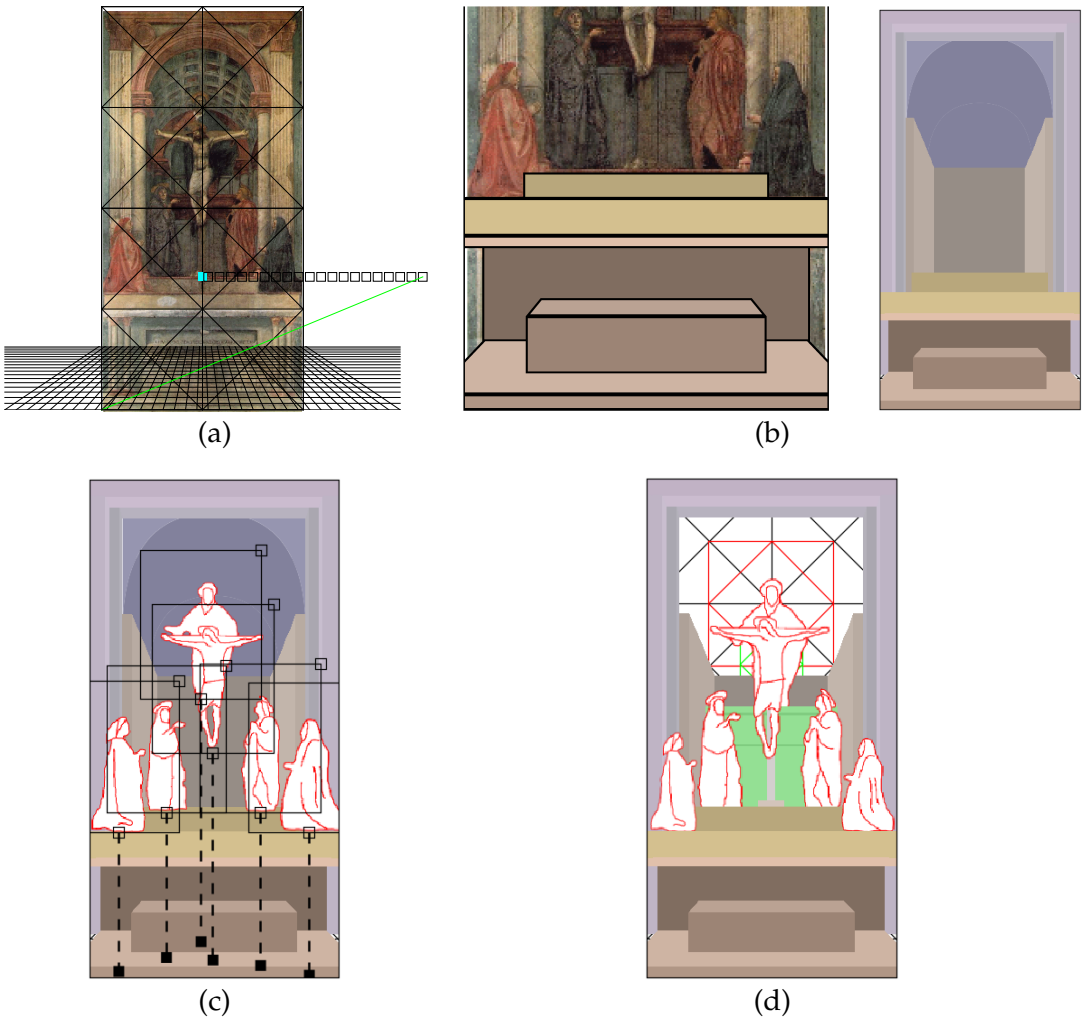
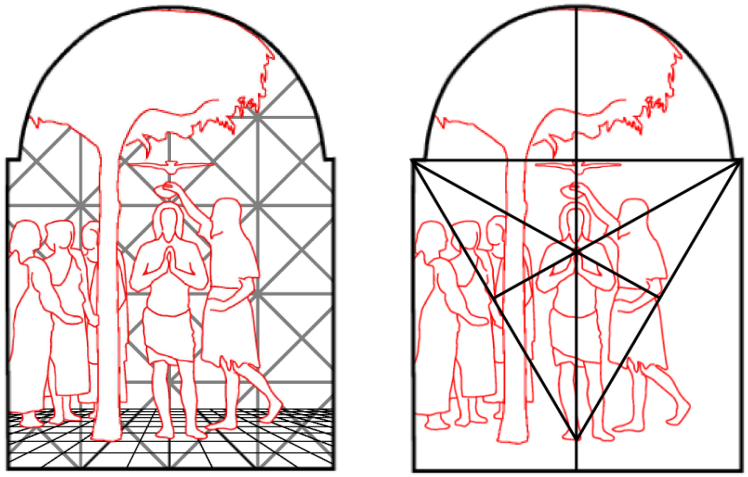


Figure 8.3: Steps to produce a schematic representation. (a) presents steps 1 to 3 completed. The pair of images in (b) illustrates step 4 during which the architectural volumes are conceived. (c) illustrates step 5, the foreground figures are roughly placed. (d) illustrates step 6, the panel are adjusted considering also alignments with the planar pattern.

8.2.1 Focus on Composition

The dominating compositional element of *The Baptism of Christ* (c. 1450) by Piero della Francesca (Figure 3.6 (a)) is an equilateral triangle (§ 3.2.3). I imitated the painting composition, Figure 8.4, by placing within the picture frame seven panels, the content of which represents foreground elements: Christ, John the Baptist, the dove, each of the three angels and the tree of life. Each is manipulated to achieve relationships of Piero's painting.

Talbot's planar pattern was the compositional reference: it assists positioning of important features (Figure 8.4 (a)). For example, most elbows are aligned along a horizontal line of the pattern. The elbows of John the Baptist and the middle angel coincide with vertical lines symmetrically flanking the central vertical line which bisects Christ. Christ's hands join at the center of a central lozenge while the dove is aligned with the central horizontal line of the lozenge above. Overall, the symbolic divisions between the earthly, human and spiritual realms are aligned with Talbot's planar pattern, which divide each figure into three parts. Diagonal relationships also exist, the baptizing arm and the bend leg of John the Baptist, for example, lying along Talbot's diagonals.



(a) (b)

Figure 8.4: The organization of the foreground figures of *The Baptism of Christ* performed with my framework. (a) shows the result with the tiled floor I used for organizing in depth and with those parts of the planar pattern that guide its composition. (b) retrospectively superimposes Carter's geometry to demonstrate equivalent correspondences.

This result demonstrates the power of the planar pattern used in my framework: geometric relationships are clearly observed. In Figure 8.4 (b), the triangle identified by Carter [69] is superimposed. The geometry of the triangle and its bisection lines coincide with important elements of the composition, as was described in Chapter 3 (§ 3.2.3). The important points of the equilateral triangle—base, apex, centroid, bisectors—lie at significant points in Talbot's sub-grid. The correspondence between my result and its relation to Carter's main geometry remind us that no single planar geometry exhausts a good composition. In fact, it is likely that a better composition reveals more planar geometries. A few directions, horizontal, vertical, 45°, 30°, 60°, and a few ratios, $1:1$, $1:2$, $1:3$, $2:3$, are pleasant to the eye, and easy to obtain.

While not all of the paintings had a visible tiled floor, for my schematic representations a tiled floor provided a spatial guide for arranging the figures in depth (as demonstrated in Figure 8.4 (a)). Christ's toes follow floor orthogonals: the central one and the one to its right. The Baptist's right foot is one orthogonal further, laterally balancing the base of the tree. Following the floor horizontals, Christ's feet are at the depth of the tree, while the Baptist is aligned with the foremost angel. The angels are positioned diagonally toward the dp to the right of the pvp. The attachment points are set to create coherent occlusions between the foreground elements, reproducing those chosen by Piero. The perspective tiled floor, even in a landscape, gives a spatial structure that organizes the arrangement in depth.

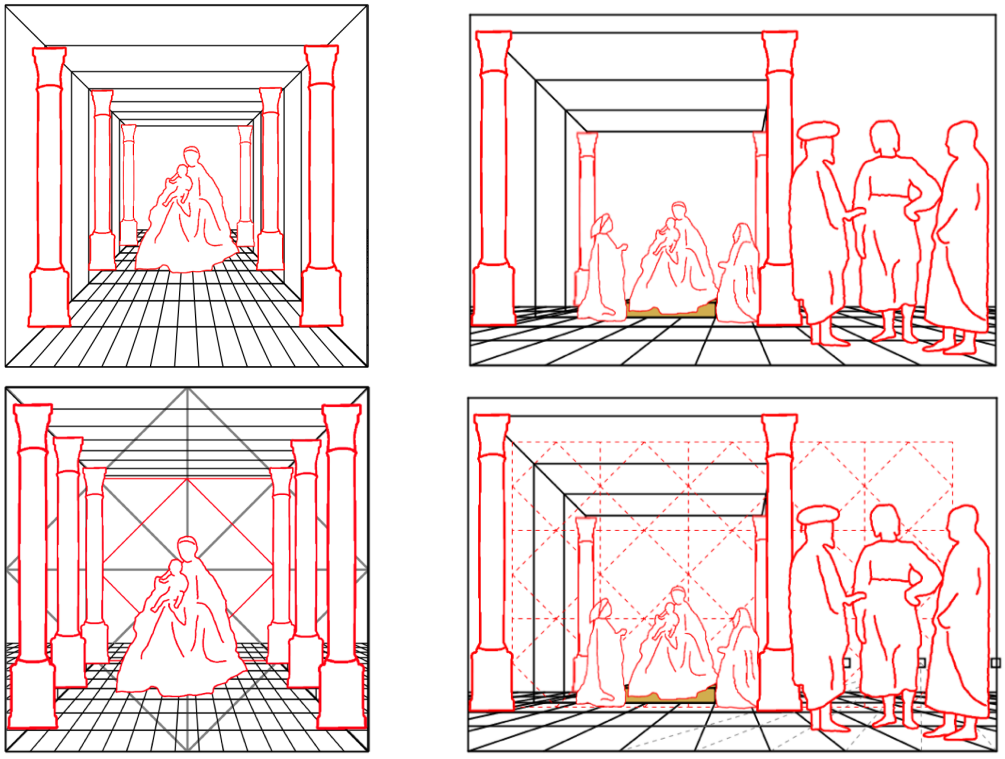
Producing the planar composition on an unruled canvas would have been more difficult. Composition is complex. There may be only a few figures but each has many salient points. Positioning them so that visual relationships prompt symbolic relationships is a depiction challenge in itself. In my framework, the tiled floor and Talbot's pattern were essential for composing in the frontal plane while providing arrangement and occlusion in depth.

The Baptism, set in a landscape, uses few elements of perspective. The next two examples, Figure 8.5, illustrate my framework when the perspective geometry is important to the composition. The tiled floor provides a sense of space, fixing the view and helping to position the foreground. Basic walls extend the foreshortening of the tiled floor to provide the feeling of a closed space.

Figure 8.5 (a), based on *Madonna and Child* (c. 1437) by Jan van Eyck, studies the effect of different perspectives on a scene; Figure 8.5 (b), based on *The Flagellation of Christ* (c. 1458) by Piero della Francesca, explores an asymmetric scene made in central perspective. Different figures are used in the background of my compositions: the Madonna and child replace the flagellated Christ; the pious women of Masaccio (Figure 8.2 (b)) the torturers.

The two different perspective configurations of Figure 8.5 (a) are different in atmosphere[1]. In the lower picture, the Madonna is centred in depth and resting in front of the base of the first level inscribed square: the dp is on the picture frame. In the upper picture the room volume is altered, on account of an alternate dp: it is behind the second column. Mary is changed to stay at the same physical location, her size changing with the perspective. This raises her in the picture to above the base of the inscribed square: the picture composition is modified. The perspective makes the Madonna remote from the viewer, her body centered at the pvp, with

[1]The perspective geometry of the painting: its architecture is known to not adhere to a single pvp. Thus, my interest to explore the effect of perspective on a room volume.



(a) (b)

Figure 8.5: Simple compositions. (a) based on *Madonna and Child* (c. 1437) by Jan van Eyck: two different perspectives of a same room volume. (b) based on the *Flagellation of Christ* (c. 1458) by Piero della Francesca.

some of her expanding robe in the upper half of the picture. In the lower picture gravity feels stronger, the overall composition is more human, with the head of the Madonna at the level of the viewer. This example demonstrates the strong influence on the composition of a small change in the perspective view. My framework welcomes such exploration.

Next, I examined a problem picture, *The Flagellation of Christ* by Piero (Figure 8.6). A full replication might be beyond the limits of my framework, but it is an opportunity to explore a more creative use of perspective. The problem is its geometry, about which many art historians have written, and among them a computer scientist [22].

While the geometry of *The Flagellation* is unusual for its time, its perspective seems conventional with the *pvp* on the vertical mid-line. However, the stability normally associated with central perspective is absent. This provokes the following questions: How does the painting get its effects? How was the geometry constructed? What can be learned from the painting geometry? There is little agreement among art historians as to the answers, so that exploring the picture's geometry is a good challenge.

In my replica I tried to capture the painting's controlled imbalance. The narrative center is in the decentered room on the left, the loggia. The three figures in the right foreground redirect attention away from the loggia. The *pvp* is uncontroversial; the *dp* remains in question, largely because it is hard to tell which picture elements are square. Feeling in the loggia for a similar atmosphere I chose a *dp* inside the picture frame, as the few dashed floor diagonals on the lower picture of Figure 8.5 (b) indicate. This practical choice gave enough depth for the figures at the back. Considering the relatively low horizon this *dp* provided more space than one on the frame; bringing it further in makes the tiles too elongated. The division of the loggia in three is an additional constraint to consider. The three subparts are marked by the four columns in the center. There are structural and visual marks on the left in place of the missing columns, the omission of which ties Pilate more tightly to the action and opens a space for the second torturer—my replica omits this detail.

My perspective selection is one of many. Piero's *dp* is difficult to estimate precisely. Using his painting as a backdrop while exploring, I could see a *dp* far out of the frame to the right, at an indeterminate distance. Choosing an appropriately distant *dp* was problematic: my tiles are limited to squares, and introducing sufficient depth to arrange the figures produced horizontals so poorly defined as to compromise the modelling task. My *dp* participates in recreating an ambivalent picture geometry feeling: brought partially within the picture it produces tension sufficient to encourage the viewer to look away from the picture *pvp* center.

Just where is the *dp* in the *Flagellation*? According to a thorough examination of the painting, assuming some likely squares in the architecture geometry, the *dp* has been estimated between 145–150 cm away from the *pvp* [45], the result corroborating with previous analyses. It is quite remarkable given the evidence, that Piero achieved a perspective that 'looks accurate but has been proved to be so' [37, p. 99], and a complex tile pattern in the loggia whose foreshortening is 'tightly compressed' and is admired for its 'great coherence and subtlety' [65, p. 30] using such a far *dp*. There is a balance when measuring among exactness, consistency and practicality using rudimentary tools.

From my exploration I am inclined to think differently: I conceive that the architecture supports many regular geometries—not necessarily square—which helped Piero. The scene geometry emphasizes line orientations other than the usual 45° diagonals: on the horizon several consistent converging points might exist, providing exactness and practicality. It could as well be that another type of convergent point on the horizon was used. Instead of using the far $dp = vp_{45^\circ}$, the point on the picture frame, vp_θ for $\tan\theta = 1/3$ might have helped the construction as the diagonals across the entire loggia volume seem to point at it; it might even correspond to the diagonals through 1×4 of the red tiles. Since many repeated regular shapes

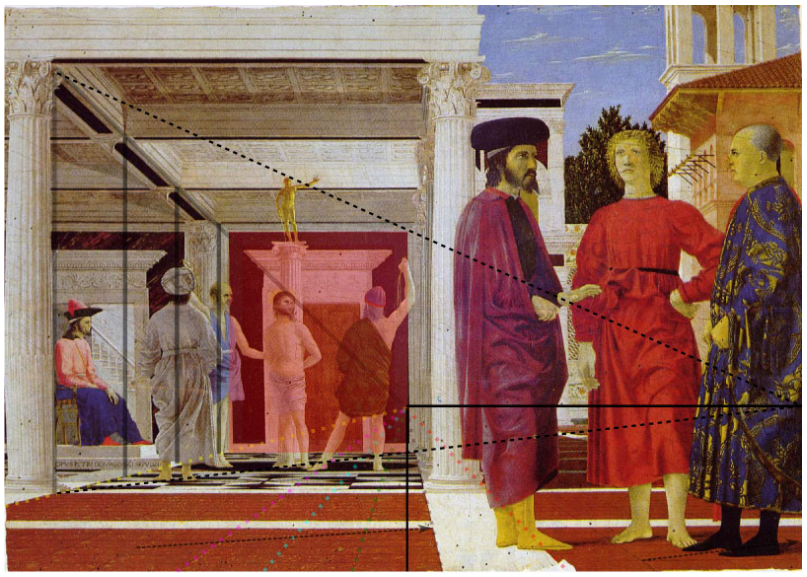


Figure 8.6: *The Flagellation of Christ* (c. 1460, 81.5 × 59 cm) by Piero della Francesca. Overlaid lines show the agreed upon *pvp*, which is on the vertical mid-line of the painting. The *dp*, well outside the painting as shown by the arrows on the floor is hard to estimate: assuming that either of the red tiles or each of the loggia divisions are square, it is at a distance between 145–150 cm of the *pvp* [45].

exist (1×3, 8×6 and 7×4 or 5) any small integer ratios are potentially practical alternatives to the usual $1 : 1$ ratio of the distance point. On Figure 8.6 I propose the horizon point at the right frame edge as one good candidate: several construction lines elaborating the architectural space point thence. My speculation agrees with authorities, art historians who believe that the painting composition and perspective follow a planar geometry on the painting surface [102, 65, Fig. 30, Fig. 43].

This painting suggests 'separate [perspective] constructions' for each the architectural space in an 'evolving design' [65, p. 30]. Piero used individual projection not only for the figures, but also for the geometric forms: 'element after element can be matched with corresponding propositions in Piero's treatise on perspective' [37, p. 101], each proposition being an elaborate perspective diagram for a piece of architecture or a figure, such as those shown in Figures 4.12 and 3.9. In effect, in the painting one can observe a boundary between two disconnected scenes, close to the vertical mid-line[2], a natural separator between the two parts of the painting. Piero demonstrated here the potential of perspective to embed one 3D world within another.

[2]In the back of the scene the torturer's raised arm curves around the vertical mid-line, the straight whip lying along it.

This observation is significant as a future avenue for research, one I did not expect as it is opposite to my initial intuition. Rather than extending the variety of geometry simulated in 3D, less regular tile shapes for example, a powerful approach to extend modelling may be creating a variety of co-existing perspective constructions, related as small integer ratios. A collage of geometric perspective constructions may be conceived using the same horizon and different vanishing points by a framework user. The viewer then appropriately follows the picture narrative by changing focus, gazing from one perspective to another. Diverse geometries can be considered in their picture existence, defined by vp_θ rather than considering their physical forms, which are limited by practicality in the finite working space of the picture.

With these compositions, which have a minimal perspective setting—a few simple lines that emerge from within the 2D pattern—I have shown that artists have a wide range of geometric expression. My framework is useful for reproducing the geometries and for understanding their organization but does not exhaust the inventiveness of Renaissance artists. In effect, working on these compositions in my framework reveals that the environment implemented is favourable for expansions as artist practices can be revealed.

The remaining examples differ in that the principal perspective is more strongly modelled: the architecture is simulated on the picture, providing a stronger perception of depth. Including the geometry of depth volumes, even in block forms, combines the effect of planar composition with arrangement in depth at another level.

8.2.2 With Architecture

Saint Jerome in his Study (c. 1510) by Vincenzo Catena shows a space unified by the architecture, and furniture of a room in central perspective (§ 3.3.1). Figure 8.7 shows my schematic representation of the painting, illustrating my framework to advantage. My composition has a calm atmosphere, like the painting. The rigid architecture of the room is schematically encoded with twenty-five blocks: in hierarchies they recreate the walls, with their wood trim decorating their plain surfaces, the placement of the benches, and the slightly elevated podium at which Saint Jerome is reading. Elements, approximated by panels, create the narration of this studious room: Saint Jerome, the dog, the bird, the pair of slippers, Christ on the cross, and the bookshelf doors semi-open along the back wall.

The panels are placed and sized so that each element takes up the desired area of the picture; an attachment point sets each element's depth in the room. But more importantly, each panel's attachment point is adjusted to produce appropriate occlusions: guided by eye

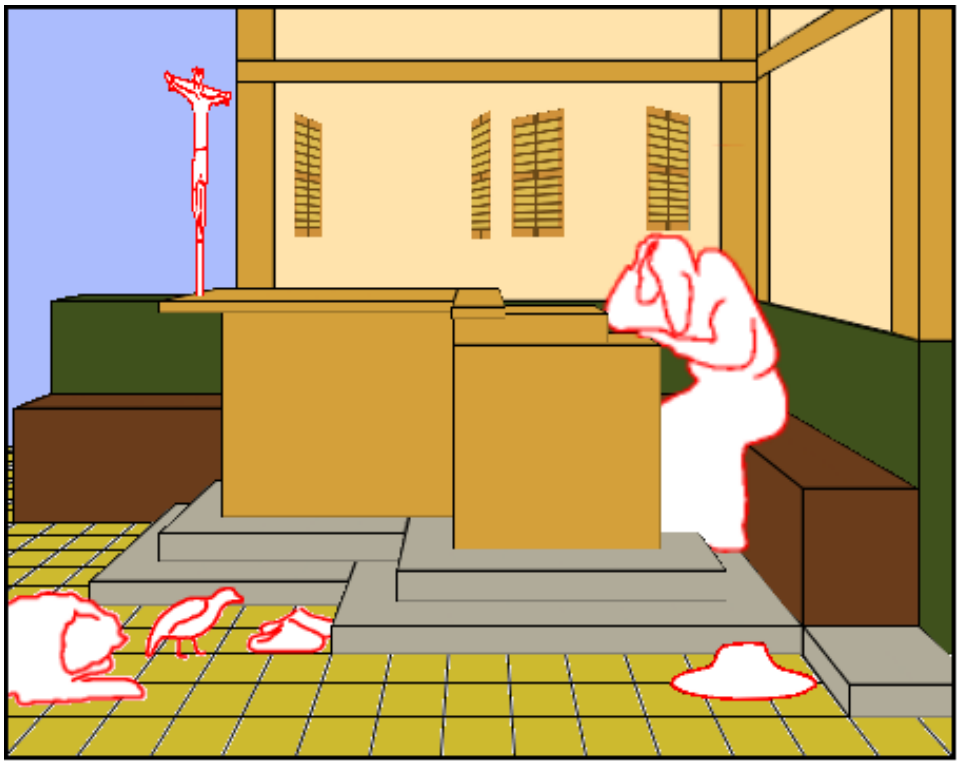


Figure 8.7: My schematic representation inspired by *Saint Jerome in his Study* (Figure 3.7 (a)).

the user's hand adjusts each panel definition among nearby shapes. Each is moved to tune the composition so as to produce the combination of occlusions and visibility that appropriately glues together the volumes and the elements of the scene.

The adjustments for Saint Jerome and the slippers demonstrate the power of the panel representation, which does not discourage interplay with nearby shapes. A pair of slippers on a panel has the pedestal occlude the toe of only one slipper. This effect is achieved with the adjustment of a single point, the attachment point of the panel. On the panel one slipper is depicted behind the other, its sole slightly higher and its toe slightly ahead. In effect, the externally created content includes together two elements at different depths. It might seem to be inappropriate content for a panel. But the attachment point provides an opportunity to make the awkward layout work. The user adjusting the panel of the pair to fit into the composition, discovers an occlusion that reinforces the weak depth cue on the panel: the occlusion feedback from the interactive rendering of my painter's algorithm, the user notices the effect[3]. The attachment point is inside the tile on which the slippers are placed. Thus, the slippers are

[3]I was not planning the effect, it was created as a happy accident.

painted before the pedestal immediately to their right. The power of the attachment point gives capabilities that lie just short of deception, the sort of coherence with which artists naturally juggle.

Similarly for Saint Jerome, manipulating his panel and its attachment point make the bottom of his robe disappear into the non-existent hole of the desk. The flexibility of the attachment point does more than grounding Saint Jerome in the scene depth: it also slides him among the architecture. The discretization of the tiles, constraining the sizes of architecture blocks plays an important role, as it specifies the granularity of occlusion.

In addition, rotated panels are used. Their content is easiest to create face on when working in the studio, as a texture is easiest to make as an orthogonal projection in imaging software. Panel rotation place them in the scene at a different orientations. There is a subtle foreshortening that reinforces the orientation. Artists use this technique manually, warping a depiction using an anamorphic grid: my framework panel rotation automates it. Partially opened bookshelf doors are simulated by rotated panels, each containing the same simple shutter texture. Once rotated they contribute lines pointing away from the pvp, disturbing the picture's central perspective. Similarly, Christ on the cross, originally created for Figure 8.2 was reused here on a rotated panel. The picture requires Christ to face not the viewer, but Saint Jerome. Thus, its panel is rotated until it is almost along Saint Jerome's line of sight, and the viewer sees the pair for the most part in profile.

This panel rotation of Christ on the cross illustrates the limitation of the manipulation in my current framework. Panel content that depicts a volumetric object cannot be convincingly rotated by a large angle [40]. The problem is that Christ's silhouette, which was represented face on, lacks inside occlusions when the panel is rotated. Catena painted Christ's body in relief; it is missing in Figure 8.7. This limitation is addressed in future work (§ 9.2.3).

Figure 8.7 emphasize three capabilities of my framework. The first is that my framework permits constructing a schematic representation that gives a strong impression of the principal perspective: the many blocks representing the room architecture and furniture produces this feeling. The second important capability demonstrated in this example is the sufficiency of the attachment point to fit an object made externally as a flat representation in the perspective picture. In particular, this example has shown how the attachment point flexibility compensates the flat representation by creating opportunistic depth cues among the nearby shapes. Finally, Figure 8.7 illustrates the use and effect of panel rotation to change the view of the panel content: it works well for representing flat content oriented differently, and less so when the content ought to have protruding aspect.

8.3 A Scene Comparison: *The Last Supper*

I now compare two schematic representations, derived from two Renaissance paintings that illustrate the same narrative: *The Last Supper* (1495–1497) by Leonardo da Vinci and *The Last Supper* (1464–1468) by Dirk Bouts. The two artists approach the scene quite differently despite the similar narratives. In this section I concentrate on the visual contrasts in planar composition and scene arrangement, which greatly affect the overall quality of the picture. Their quality is related to familiarity with perspective. da Vinci's perspective is mastered as a technique for realism that enhances composition; whereas Bouts' painting showcases his perspective ability. This example illustrates that restricting the perspective view and the development space to the canvas surface is not a guarantee of good composition.

Figure 8.8 shows my two schematic representations with the planar pattern apparent (above are the actual paintings also with the pattern, so that the compositions can be compared). Both have Christ in an enclosed room surrounded by his disciples, taking the last supper. My reconstruction shows Mary around whom nine figures are gathered.

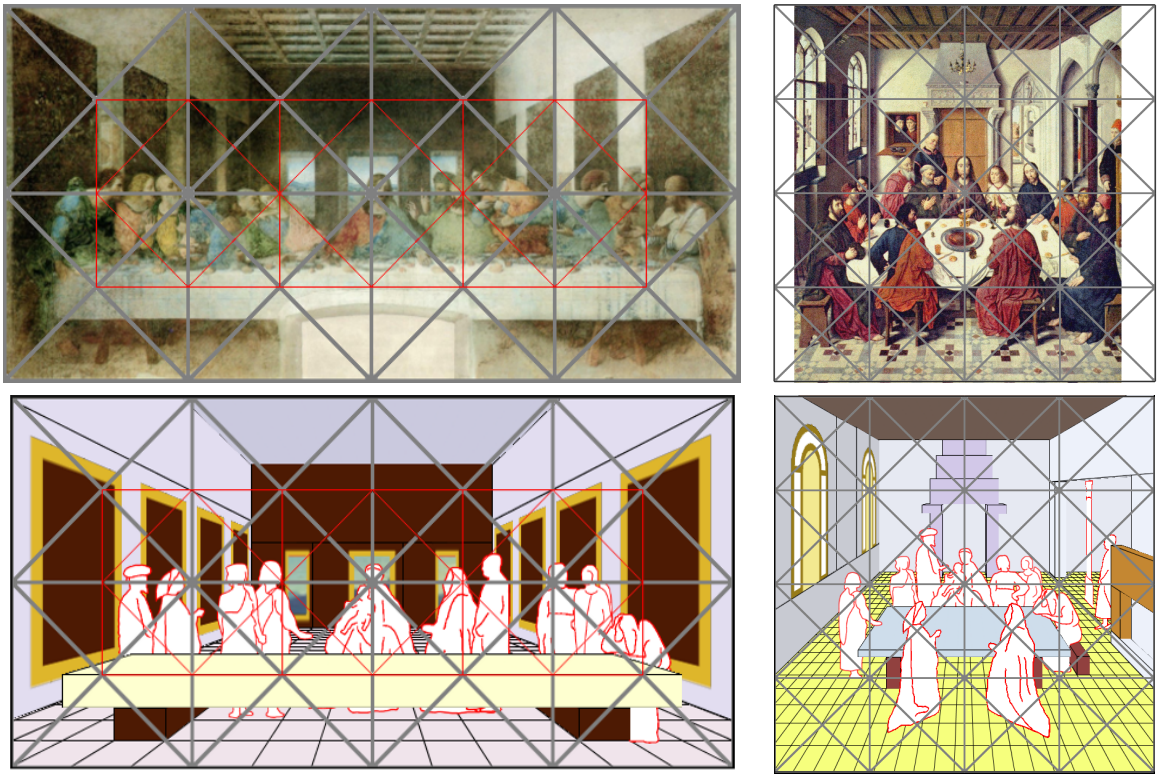
8.3.1 Centers

In both paintings, Figure 8.8, the *pvp* is on the vertical mid-line of the picture. However, the horizon is not similarly placed, giving different views of the scene: da Vinci's horizon places the viewpoint at the height of Christ's head; Bouts' horizon is much higher, so the viewer hovers above the scene.

Thus, my replica of da Vinci, Figure 8.8 (a), uses a *pvp* at the geometric and physical center of the composition (a balancing center [6, p. 225]), which Mary's head hides. The surrounding protagonists' heads are therefore along the horizon, varying only by individual pose.

In my replica of Bouts, Figure 8.8 (b), the high horizon crosses the lintel on the back wall. It makes the table top visible, exposing on the mid-line of the painting the circular plate with the lamb roast. Bouts uses more centers than da Vinci: the *pvp* for the 3D arrangement, Christ's raised open palm for the 2D composition, and the table for the narrative. Bouts' high viewpoint places the figures' heads at different heights on the picture: they rise toward the horizon as the figures recede to the back of the scene, despite similar physical heights and poses.

The *dp* represents the station point, a center, from which the viewer observes the picture. For my schematic representation I chose a *dp* that creates a feeling of room volume similar to the original. Exploring the possibilities in my framework, I found that placing the *dp* on the



(a) (b)

Figure 8.8: Schematic representations of *The Last Supper*, below the paintings that inspired them: (a) is Leonardo da Vinci's painting (c. 1497); (b) is Dirk Bouts' one (c. 1468).

frame was natural for the da Vinci Figure 8.8 (a). In contrast, in Bouts' Figure 8.8 (b) the dp is hard to determine. The orthogonals on the decorated tiled floor converge to the pvp, as do the side windows, but the floor diagonals point towards a dp far outside the frame, which the rest of the scene construction does not support. A dp outside the picture makes the space feel shallow, but in the painting the room feels deep and the supper scene impression does not seem distant. Thus, I considered the tiled floor dp misleading and chose instead the traditional dp: at the picture edge[4] the dp provides the appropriate depth and room volume.

8.3.2 Room Perspective

The perspective view of the room architecture gives each painting a different overall look and feel. The two configurations vary most in the empty area around the table, which provides space for the protagonists to live their narratives.

[4]I consider the dp at my picture frame that is square. The awkward proportions of Bouts' painting makes it doubtful that it preserved its original dimensions.

Leonardo's perspective configuration is open, its depth accessible: the back wall is not close and has windows onto a far horizon. The room is deep, without voids in the picture space and gives a strong sense of 3D. In addition, the wide rectangular frame gives lateral space to accommodate a spacious room holding the many guests. As indicated by Talbot, the painting's dimensions and its perspective fit his planar pattern [102]. Figure 8.8 (a) uses a two wide one high combination of the basic pattern for the picture frame. I enhanced alignments in my schematic representation, such as making the vertical mid-line of each lateral wall coincide with a pattern line. My framework's construction of the room creates a convincing 3D volume where the architecture is balanced and well proportioned.

Bouts' perspective and picture frame dimensions spread the composition in depth and in height, making it difficult to use the canvas space effectively in my replica (Figure 8.8 (b)). The focus of the narrative is located well below the *pvp*, the picture frame has awkward proportions, that do not correspond to the planar pattern used in my framework. I used a square frame, which (re)-adds parts of the room symmetry on the left wall.

While defining the room volume of Figure 8.8 (b) I concentrated on the walls, the back wall with its prominent architectural structure, the left wall with its pair of windows, and the central pillar of the ceiling an imperfect guide that I tried to maintain. The table surface I made an effort to match. All the same my imitation of the room volume and its furniture are awkward and far from matching the painting. (My table surface is more stable; it gives less of an impression that an invisible supper is about to slide.)

My visible tiled floors, determined by the locations of the *pvp* and *dp* discussed above, make explicit the basic arrangement in space compared to the originals, where the floors' tiles are either hidden or limited to the front of the scene, not to say inconsistent ('the bare floor of the foreground' [6, p. 185]). In Figure 8.8 (a), my tiled floor accords with the harmony of the scene architecture. The tiles look square, well proportioned and balanced. In contrast, in Figure 8.8 (b) the high horizon forces a slanted floor, with tiles elongated in depth. In effect, the floor tiles further emphasize misdirections of Bout's painting. For Arnheim, 'in spite of [Bouts' *Last Supper*'s] otherwise realistic style, some of the flat frontality of medial icons survive' and 'the upright format of the picture stresses its Gothic, vectorial verticality' [6, p. 185].

Within the room's perspective foreground figures must be positioned. Counter-intuitively, architectural configurations that leave significant areas for foreground figures are a sign of misfortune to follow. Bouts' perspective is difficult to match and the foreground figures will not be easy to place. da Vinci's perspective space emphasizes composition; it is as much simpler to reproduce as it is realistically better for both 3D simulation and 2D composition.

8.3.3 Composition of the Figures

In the disposition of the figures, the approaches of Bouts and da Vinci differ: Bouts emphasizes realism, da Vinci composition. In Bouts' painting, the positions of the figures around the table seem documented, not composed, as in a news photograph; da Vinci transgresses against 'realism'[5] for the sake of composition.

Achieving Bouts' composition required many 3D manipulations, some awkward, to achieve reasonable locations and sizings. Owing to his peculiar perspective I put much effort into the figures, systematic cascading down sizing relations, yet achieved unrealistic results. My framework features helped, but gaps having both 2D and 3D salience were hard to fill. Overall the space left by the furniture did not fill naturally with figures.

Given the architectural layout of Figure 8.8 (b), occupying all four sides of the table with diners is a requirement. Some panels were placed on each side, but all sides are not of equal quality: five figures fill the far end of the table with gaps and overlaps among them properly managed; the lateral sides of the table are inadequately empty; the nearest table side presents figures in bad positions in almost back facing poses. When Bouts turned the heads of the two disciples, I used panel rotation to adapt the figure orientations, redirecting them towards the narrative center of the picture.

For consistency Bouts' scene requires foreshortened figures. Some sizings were adjusted by eye, but I also relied on my framework to move elements in depth. Specifically, the two closest figures, slightly rotated, were sized in reference to the man standing behind Mary: aligned horizontally they were brought closer while they resized automatically (§ 5.4.2). The planar pattern supported symmetrical sizings and placements: the rotated figures are at opposing grey intersections; a vertical line separates figures on opposite sides of the table.

In contrast, da Vinci's figures are on stage, with the viewer in the audience. For Figure 8.8 (a) the composition was assembled using local adjustments of the nine panels. da Vinci's geometry was effective, as the schematic representation demonstrates. In effect, placing and sizing were easy: the horizon, the tiled floor and Talbot's pattern providing guidance. With fewer figures, the side and ends of the table gather them together. Grouping gives a sense of engagement in the event: the planar pattern lines separate groups, with the symmetry around Mary's vertical axis disturbed to add dynamism. In each group, figures lean toward one other, using simple occlusion among them. Moving slightly an attachment point interchanges the occlusion within a group. Only sizing the seated Saint Jerome required compromise: his partial occlusion by the table must be coherent, while his proportions consistent with the other figures.

[5] In this instance it is worth pointing that what is more realistic depends of culture and context.

The picture has a balanced 2D composition, the figures disposed symmetrically around the principal viewpoint, the site of the main protagonist. The staging is more important than the 3D realism of the representation. For da Vinci, the ideal viewpoint is the position around which the scene is planned and presented. Picture elements should seek positions and orientations that show them at their best. The viewer should be able to see all regardless of 3D realism. Thus, for example, back facing positions are avoided.

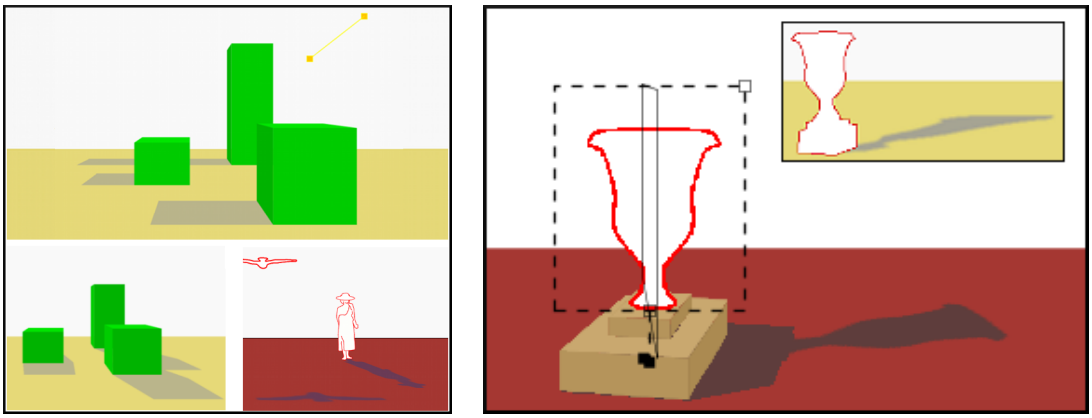
Reproducing each painting made it obvious that the viewpoint and picture frame chosen by da Vinci are particularly well adapted to the spiritual narrative compared to the more makeshifts choices of Bouts, giving it a 'mundane setting' [6, p. 185]. The schematic representation based on da Vinci is both more pleasing and simpler to arrange. The *pvp* and of the picture size choices contribute to the harmonious environment in which the picture's narrative is expressed successfully. The 2D composition having been decided first, the 3D architecture in accord around it, is effective for both the picture quality and realism. Contrary, fitting a 2D composition in a space that is primary 3D is difficult.

Overall, this comparison demonstrates the significance of the location of the *pvp* within the picture frame. The *pvp* creates a strong context determining the realization of the composition. This effect is mediated by reflective flows among the perspective, the space, the architecture and the composition. Compared to 3D modelling software my framework makes it possible to adjust the perspective while positioning picture elements within the picture as shown in Figure 8.5 (a). In 3D modelling software the camera has six degrees of freedom, the image plane and frame four more. My framework strongly restricts the viewpoint with respect to the picture plane, allowing only restricted motion: a few discrete values laterally and in depth and a continuous value for height; the line of sight perpendicular to the picture.

Despite the viewpoint constraints, *The Last Supper* comparison demonstrates that 2D composition within a 3D simulated depth on the picture does not guarantee good composition. The two configurations reproduced in schematic representations show that my framework helps a user to learn the geometry of a picture.

8.4 Projected Shadows

Projected shadows are included in my framework to show how adaptations to the construction and the algebra encompass a more complex geometry and because they contribute to a perceived depth in compositions. Projected shadows can reinforce the shape of an object—despite



(a) (b)

Figure 8.9: (a) Pictures with projected shadows that indicate different illuminations. (b) My experimentations to reproduce Vanderlyn's shadow appearing in the painting (Figure 2.2 (b)).

it being merely a panel—with the surprising effect of augmenting the object. Adding projected shadows to a schematic representation can provide depth cues, often acting as approximate attachment points for off-the-ground objects. The results of this section demonstrate the spatial function and attributes of shadows.

Shadows are rarely foci of a composition, and the sharp edges of projective shadows give them too much prominence for most compositions[6]. Yet the inconspicuous shadows in *The Astronomer* by Vermeer show the projected shadows are the right place to start, softening the edges by eye later (Figure 1.1 (b)). Thus, Figures 8.9, 8.10 and 8.11 should be taken as proofs of concept demonstrating the spatial contributions of projected shadows.

Figure 8.9 (a) shows several types of illumination casting shadows on the ground plane. The upper picture's shadows shows mid-afternoon illumination coming from the right. The absence of foreshortening indicates a distant light, or in computer graphics' terms a directional light. (The oblique line indicates the direction.) While the shadow effect is dull, it solidifies the blocks anchored on the floor. Below it, a similar scene has shadows lit by a light deep in the scene, on the left. Here, each block's shadow is cast by multiple faces and three dimensionality of the blocks is reinforced.

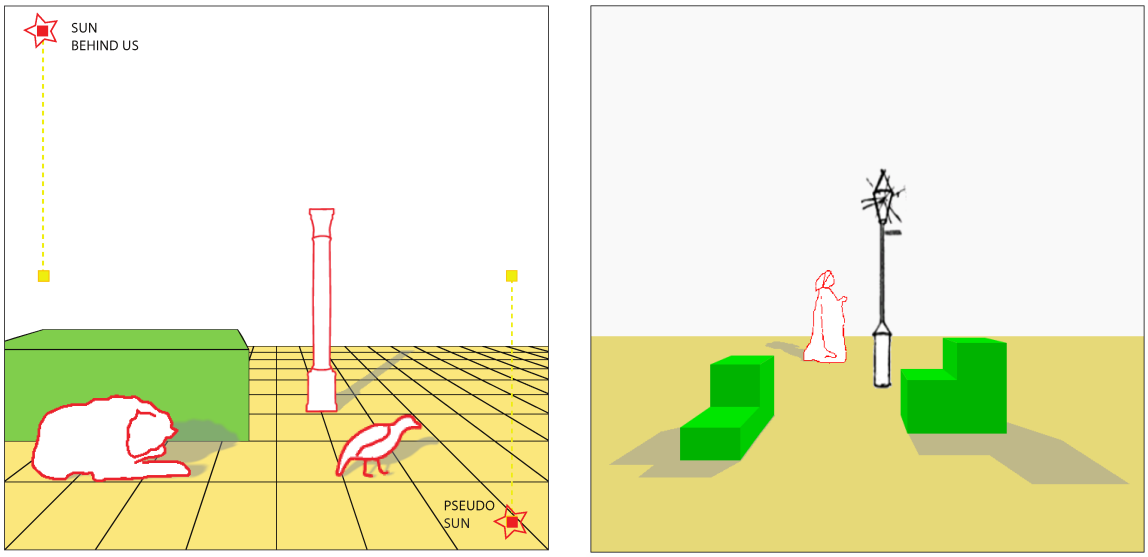
The last picture of Figure 8.9 (a) illustrates projected shadows cast by the contents of panel approximating foreground elements. The dove from *The Baptism of Christ* by Piero della Francesca (Figure 3.6) creates another reality when cast on the ground plane, as it in our habit to see birds much in their shadow forms on that plane. In addition, a hat is inserted on a

[6]De Chirico is a modern painter whose paintings have many hard-edge shadows that reinforce the tension-filled inconsistent perspectives of his paintings.

third panel and placed on the woman's head. The projected shadow shows no gap because the hat's attachment point stipulates the depth of the floating object. (This picture uses simple parameters that modify the appearance of shadows, as described in Appendix D, § D.4.2.)

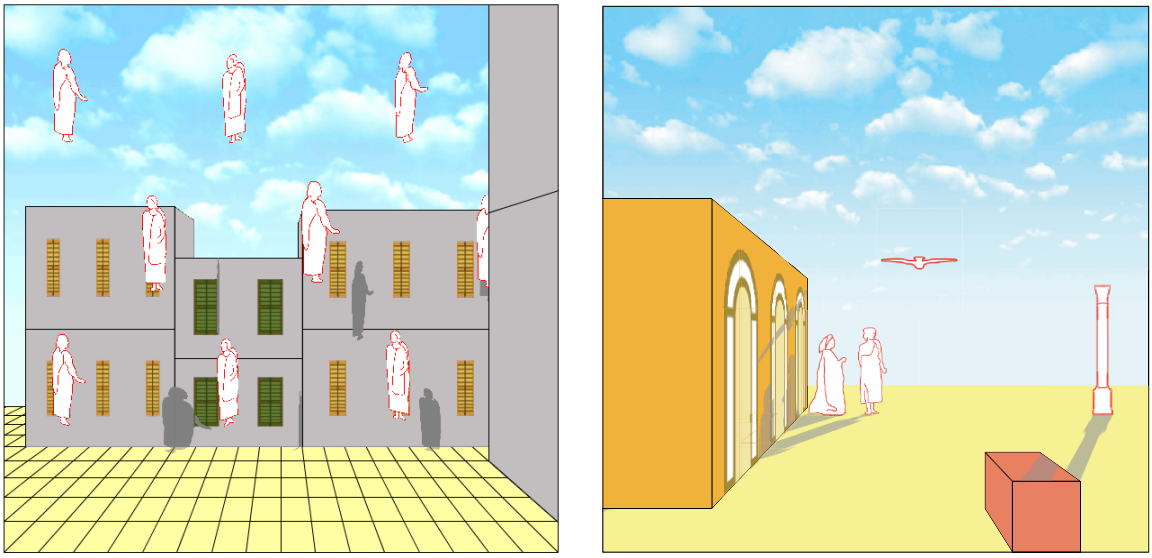
Figure 8.9 (b) presents a slight modification to reproduce the shadow cast by the vase in Vanderlyn's painting (Figure 2.2 (b)). As shown in the insert picture, a panel whose content is the silhouette of the large vase casts an unsatisfactory shadow onto a ground plane with respect to a pseudo-light. However, the main picture shows an improved version created with my framework. As my framework handles rotated geometry, the shadow instead uses a rotated version of the panel, invisible in the final picture: a kind of billboard with respect to the light. The boundary of the panel casting the shadow is shown in Figure 8.9 (b) to visualize the geometry in play. The result is closer to the effect depicted in the painting: the shape of the physical vase is better apprehended owing to its projected shadow, as is the light location.

The pictures of Figure 8.10 include shadow effects that add an appealing aesthetic to a composition. Figure 8.10 (a) uses a pseudo-light: three panels cast shadows, with the column's shadow indicating how they point towards the light vanishing point. The (shadowless) block obstructs a ground shadow that continues on its front face, reinforcing the block's three dimensionality. Figure 8.10 (b) is lit from the centre of the picture: shadows diverge in all directions. The light is at the bulb of the street light, which is a panel set to not create a shadow.



(a) (b)

Figure 8.10: (a) a pseudo-light illuminating casting three panel shadows. (b) a lamp post creating shadows in different directions.



(a) (b)

Figure 8.11: Compositions demonstrating interactions between volumes, panels and shadows. (a) is inspired by *Golconda* (1953) of René Magritte.

Finally, Figure 8.11 has two compositions that include shadow effects. My schematic representation of Figure 8.11 (a) is a composition after Magritte in which shadows make the eye revisit the foreground elements: raining angels in varied orientations, casting shadows on building facades from a sun placed behind the viewer on the left. The shadows situate the raining angels in depth. On the picture, the angels are located close to a regular grid using Talbot's planar pattern indications. Shadows of widely rotated angels are thinner than angels that are almost face on. Figure 8.11 (b) shows shadows created by panels interacting with volumes.

While my framework has shadow capabilities, shadows must be used sparingly within a composition because their presence risks stealing attention from the remainder of the composition. In effect, shadows should be present only when they add to the narrative, which is rare; strong shadows without significance seem out of place.

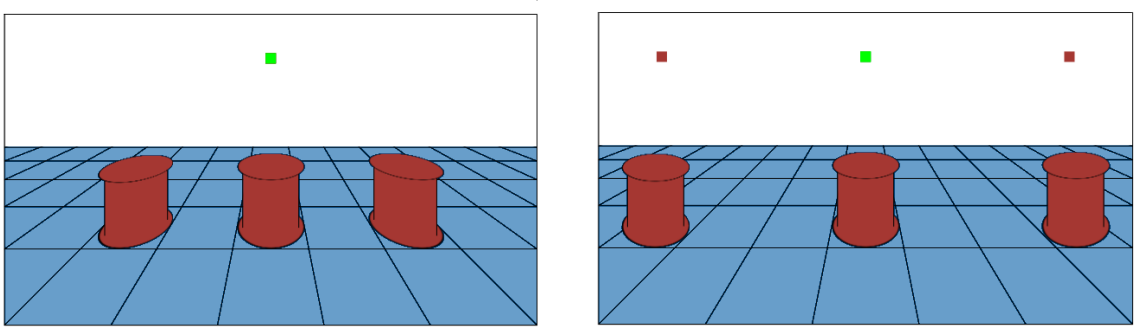
By including shadows, I demonstrated that the panel abstraction contains the necessary information for an effect often considered as a 3D one. The geometry for shadow is significantly more complex than the rest of the 3D simulation created by my framework: two centers of projection are used creating interactions with different types of surfaces, as the pictures above illustrate. The constructions I found for shadows were schematic, detailed enough to improvise any particular shadow, but far from the complete descriptions needed for computation. Thus,

I completed and implemented the constructive geometry, and formalized the constructions in algebra. The results show that the operations used when simulating 3D within 2D are robust. They can be combined to significant complexity without breaking.

8.5 Multiple Perspectives

Another capability of my framework places several perspectives in one composition by applying local perspective to volumes, which is complementary to the off-line local perspectives of panel content. The framework is designed to play with constructions. For example, Figure 8.12 (a) shows three columns, each constructed according to the principal perspective of the picture: the two lateral columns look physically different than the central one. Only the relation of their bases to floor tiles hints that they are distorted by perspective. The three columns can be projected so that they convincingly look the same: Figure 8.12 (b) shows them constructed with my framework using an individual vanishing points. They look much better and would be perfect if the tiled floor was hidden. My framework affords manipulating construction lines with an individual perspective.

Despite the constraint of working on the surface of the image plane, exploring the interplay between 2D composition and 3D appearance is easy to accomplish in my framework, which automates the computation while leaving control of the appearance to the user. Manipulating local vanishing points simulate the playful perspective exploration of more contemporary paintings, without requiring 3D modelling and the inconsistency of viewing from multiple locations. Adjusting appearance using 2D manipulations and definitions is satisfying because it reproduces better the practice of artists.

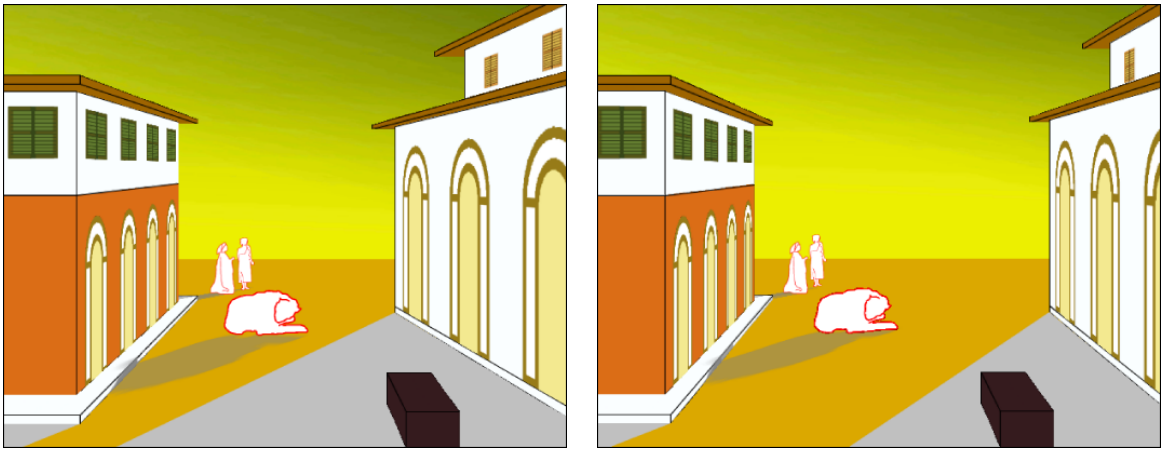


(a) (b)

Figure 8.12: (a) Columns adhering to a single *pvp*. (b) A local perspective constructs each column, using a *vp* vertically aligned to their center.

Figure 8.13: *Piazza d'Italia* (1913) by Giorgio de Chirico, © Art Gallery of Ontario, Toronto. The construction lines to the horizon defining the buildings on each side of Piazza are drawn.

Figure 8.14 (a), created in my framework, was inspired by Giorgio de Chirico's *Piazza d'Italia* (1913), Figure 8.13. I used two local perspectives to construct the blocks that reproduce the two lateral buildings. De Chirico paintings are ideal for exploring flexible perspective: they are nominally realistic, while pushing the limits of geometry. What is hard about multi-projection rendering [2] is the treatment to be given to the places where different perspectives have to be reconciled. The convenience of De Chirico is that the areas in which mismatches might happen are empty. The investigation of De Chirico's geometry using my framework is different: his composition is imitated using 2D canvas constructions, which is more in the style of De Chirico himself, who likely worked directly on the canvas, rather than postulating a 3D scene.



(a) (b)

Figure 8.14: In (a) the schematic representation follows De Chirico use of a local vanishing point for each of the building on either side of the Piazza. In (b) both adhere to the *pvp*.

In De Chirico's painting one observes that the buildings are drawn using construction lines that meet on different horizons. Their parallel lines drawn over Figure 8.13 converge above the horizon—which is set by the back of the scene where the volcano and the train stand—separated vertically and laterally. Neither building's dorsal face converge at the vertical center of the picture.

Figure 8.14 (a) was created to reproduce the main aspect of the construction analysis just described. Figure 8.14 (b) is the starting scene configuration with both buildings aligned with the picture principal vanishing point. It is a harmonious calm composition, with simple geometry. The two buildings are at rest. Figure 8.14 (a) is created by pulling the local vanishing point of each building away from the pvp. The framework skews the buildings accordingly. Figure 8.14 (a) has the same geometry as the De Chirico painting, Figure 8.13. The two have a similar effect on the viewer: the space is more enclosed; the buildings exert pressure on each other producing an anxious unstable space; the small figures in the back contrast with the middle panel, enhancing to the heavy atmosphere. In effect, De Chirico 'painted 'architecture' [..] not as it appears; it is deliberately enigmatic [..] both familiar and uncanny' [78].

8.6 Discussion

To close this chapter, let me briefly discuss two points: the basis of the features of my framework and the limitations on the evaluation presented.

8.6.1 Lessons

My framework provides features for creating and analysing the geometry governing a picture, focusing immediately on the visual space to be explored. The framework shows that restricting is important. There are three big ideas to take from this chapter.

1. The tiled floor is a tool for organizing space and simulating 3D. Its appearance, constraining how the picture is viewed, allows the user to render depth without sacrificing composition.
2. Alignments are significant in 2D.
   (a) Objects are aligned to simple geometric shapes with pictorial meaning.
   (b) Relative dimensions often reduce to ratios with small integer numerators and denominators.

(c) The *pvp* must be aligned with the 2D composition geometry.

3. Panels play the role of cartoons and are enhanced by computation.

   (a) Panels assist laying out in 2D the geometry of 3D objects. As rough sketches, portraying independently the 2D projections of figures, they are inserted later in a picture to complete and refine its geometry.

   (b) Panels are placed on the painting surface with simple manipulations to fit them within the simulated 3D geometry: the attachment point is a powerful visual cue rooting them them in depth using only simple 2D interactions.

8.6.2 Evaluation

This thesis argues that composition is essential to any effective picture and that good composition is a challenge for which no simple recipe exist. In this chapter I have shown by example that there is potentials for graphics packages to address better the user's composition task. Features in my framework help to develop a 2D composition, while the 3D simulated geometry gives an overall coherence to the picture. In addition, my framework helps a user to understand geometries used by artists: the user learns as he manipulates, binding visual appearance with specific actions.

However, my framework is only barely evaluated. To be sure, the compositions I created using my framework are at most an existential proof: there exists a user (the author) who can use the framework to create and understand compositions with simulated depth. In contrast, formal evaluations are valuable if they can be universally quantified: any user can..., or any user with artistic training can... Experiments purpose is to answer this type of questions. In a controlled environment, participants, who are typical members of the universe, are asked to perform a task and their performance is measured. To validate my framework an interesting task could be to perform the schematic representation of *The Trinity* painting, the exercise presented in Section 8.2. But, there are complications to run and analyze such experiment: the task is broad and co-founding factors are not isolated; and any performance metric biased.

That is why I plan to pursue the validation of my thesis work in three stages. The natural start is to get informal free-form feedback on the framework from artists. They are the experts on picture creation, thus listening and watching them on my framework I expect better to understand their practices and creative considerations. This pilot phase of evaluation might help to improve the framework and understand possible users.

Only thereafter, usabilities studies to investigate within a controlled environment the aspects of my framework that make it practical and efficient to use, are worth undertaking. Specific tasks should be designed. Initial experiments need to be simple and focus on a single feature, measuring the time and possibly the quality benefits of having it available. (This latter measure is a challenge as under many circumstances likely to be ill-defined.) Further experiments would test the framework as a whole, still giving the user's precise objectives to accomplish. However, enticing a framework user into free creation is the most interesting to investigate. Only with a creative task can the capability of the expressive potential of my framework could be understood. Will users who are not the implementer be creative with the tools the framework provides? There comes the difficulty to assess and judge the results.

Finally, owing to my personal interest in picture understanding, I plan to pursue perceptual experiments on different aspects of my framework, to understand when and why they work or do not work. For example, knowing whether and when the shadow approximation used in my framework is sufficient would be valuable. Using a tachistoscopic presentation of a pair of stimuli, one using my shadow approximation the other using shadow mapping, it could be determine cases when a difference is perceived.

Chapter 9

Contributions & Future Work

In its fifty year history, computer graphics research has developed the physics of light and matter into a computational structure capable of modelling and rendering scenes of astounding complexity. This achievement depends on the conceptual similarity between physics and numerical computation. The equations of physics may be complex, but to a computer scientist they belong to the familiar world of algebra. Yet physicists are not alone in their understanding light; visual artists have also created a large body of knowledge and craft addressed to light and appearance.

Prima facie, we might wonder why graphics research focussed so closely on physics. After all, the goal of the artist—creating pictures that evoke thoughts and emotions in human viewers—is much closer to the goal of computer graphics than is the physicist's goal—understanding the laws of nature. Surely, a different computer graphics would have developed had expressivity been as important as fidelity to optical physics. Or to paraphrase, had Piero conceived a computer and been capable of building it, what sort of computer graphics would he have created? This question, which is much too large to be exhausted in a single thesis, and the alternate science that would have ensued is the larger issue to which my research contributes.

In this chapter I first review the contributions of this thesis, then turn to the future, suggesting ways in which my work can be extended and complemented.

9.1 Contributions

Two ideas from the history of perspective in art animate this thesis. The first idea is the negation of a widely held opinion that the 2D composition of medieval art gave way to 3D composition in the face of perspective, then returned to 2D with the advent of the perfect perspective of photography. On the contrary, art historians observe a continuous development of 2D composition from antiquity to the present. The second idea is more recent: Renaissance artists did not use a single, unitary perspective, but freely mixed geometrically inconsistent perspectives in a single painting, giving each object the perspective that best showed its nature, albeit within an overall perspective framework.

Taken together, these observations show the primacy to the artist of two dimensionality: perspective is added to a 2D composition to bring out otherwise invisible features of the picture elements. To the artist 3D space lacks primacy. It is not the truth, of which the 2D work of art is a limited imitation. The conceptual approach of 3D computer graphics could hardly be more different: the 3D scene is primary, and its projection onto a 2D manifold an impoverished view of a more information-rich reality. Renaissance artists were not lacking the geometric knowledge necessary to give primacy to the 3D scene, but chose not to do so. The research in this thesis is thus a contribution to an intellectual theme in the conceptual foundations of computer graphics: what is an appropriate relationship between images produced by computer graphics and images produced by painting or photography?

My research makes its contribution by exploring how the methods of Renaissance painters could be supported by a computer, a prosaic way of asking how they would have pursued the geometry of image composition had they been able to create a computer. In concrete terms, the first contribution of the thesis is the panel concept. A panel has a well-defined 2D geometry: it is a rectangular surface with an attachment point where its central axis intersects the perspective tiled floor. Panels contain 2D content, usually projected versions of scene elements having perspective consistent with but not the same as the overall perspective. Panels are, in essence, a computer version of the cartoons from which Renaissance artists assembled their paintings.

Panels alone produce only a collage. In Renaissance painting the objects they represent are located both on the image and in a volume. The second contribution of this thesis is the computer implementation of volumes defined by tiled floors. The tiled floor is important perceptually—the depth of an object standing on a tiled floor is perceived directly by the viewer—and also geometrically—vertical projections of the object to the floor uniquely define its depth. Thus the tiled floor is ubiquitous in Renaissance painting, often explicitly visible, sometimes as a hidden computational device.

Renaissance artists constructed tiled floors using combinations of construction lines: tile boundaries are placed by drawing lines between points, and using specific intersections as fixed points, in the style of Euclid and Pappus. Construction lines accomplish geometrically computations now performed algebraically. The constructions handed down to us are incomplete, and rely on the artist's knowledge of the world to resolve ambiguities. The third contribution of this thesis is to show that the constructions can be made algorithmic: appropriately formalized, they are independent of artistic intuition. This is true both for tiled floor constructions and for constructions used to place volumes, objects and shadows in the scene.

The geometric calculations artists do when they draw lines on the image surface are purely two dimensional: the results they provide can easily be calculated using two dimensional projective geometry. Thus, creating a two dimensional image with correct 3D linear perspective is possible using only 2D computation. The fourth contribution demonstrates this counter-intuitive result and provides matrices representing 2D projective transformations that provide appropriate perspective calculations. A corollary of the fourth contribution, implemented but not explored in detail, is a new method for positioning objects in 3D using a 2D display, which promises to ease one of the most effortful of tasks in 3D modelling, direct manipulation of objects in 3D using only 2D input and output.

Shadows rendered in perspective are a double projection problem. A shadow appears when the silhouette of a shadowing object is projected onto an illuminated plane with a light source as the centre of projection. The fifth contribution is making algorithmic a collection of artists' constructions for shadows. Shadows interested Renaissance artists who saw them as flattened objects, similar to the flattened objects that appear in a picture. Artists have created 2D geometric constructions that place shadows correctly on arbitrary plane surfaces, even when the light source is outside the canvas. But to handle poorly defined cases the constructions require significant geometric improvisation from the artist. An algorithmic equivalent must make that tacit knowledge explicit. Doing so is one part of this contribution. The other extends the fourth contribution, which shows that affine geometry in 3D can be calculated in 2D using projective geometry.

Artist and art historian Richard Talbot created a pattern to explain the composition of Renaissance paintings, including several systematic departures from true perspective that he has observed. Talbot convincingly claims that, when using patterns like his for composing in the image plane, artists recognized sets of lines that occurred in primitive perspective drawing, thereby learning how to simulate 3D within 2D. The sixth contribution embeds the five contributions described above within Talbot's pattern, unifying the planar geometry artists use in composition with geometry that allows them to give objects apparent depth.

The perspective geometry comes first in planning a painting, in contrast to computer graphics, where the view and lighting are chosen relatively late. The geometry provides an extension of the Painter's algorithm that renders pre-ordered parts of the canvas, in a content-independent fashion while achieving correct occlusion. This, my seventh contribution, seems to depart from the practice of artists, who organize the picture to paint each area only once. However, my algorithm makes available many painting sequences, and I conjecture that among them are orders, probably content-dependent, that are much closer to artists' methods.

To test the ideas, I implemented a prototype environment for creating and analysing composition. It supports a user composing a painting from volumes, sketches, lights and shadows within an enclosed architectural volume or in an open space. It offers both direct geometrical calculation of artists' constructions and the more efficient matrix calculations, a facility that was used for comparison when exploring the correctness of the two calculation methods. I validated with my framework that it is possible to integrate into computer graphics software 2D methods that simulate the 3D geometry as it has been developed for centuries by artists, on the picture surface.

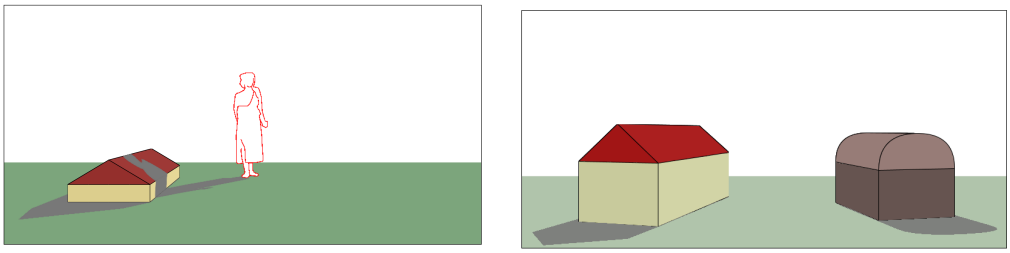
9.2 Future Work

My research considers what may seem an academic question: to what extent can 3D computer graphics be recreated using only the 2D 'compass and straightedge' geometry that was the foundation of pictorial composition? But, as is commonly the case, the answers to academic questions introduce practices and concepts that could advance the state-of-art in practical areas not considered during the research.

Sections 9.2.1 to 9.2.3 discuss concrete and immediate enhancements for my framework that could lead to deeper avenues. Section 9.2.4 contains more speculative ideas.

9.2.1 More Geometric Shapes

Blocks and panels are adequate primitives for a proof of concept. Artists and stage designers, after all, use them as proxies for volumes and images when they rough out a composition. Panels are rich enough to encompass cartoons, the artist's preferred way of introducing independently-generated content into a composition. Both the volumetric constructions and the painter's algorithm are based on blocks. But blocks are too primitive to stand for the architecture and furniture of a worked out design.



(a) (b)

Figure 9.1: Different volumes constructed by my framework using 2D computations: the roof-like shape uses three planes; the semi-dome uses interpolation to simulate a curved surface.

More complex geometric volumes can be local enhancements of blocks, with each new shape contained within a block. Other refinements of architectural volumes can be achieved by considering planes other than the ones used for a block: the octagonal prism of Appendix F and the roofs of Figure 9.1 are two examples. Drawing non-planar surfaces using interpolation is also effective: the columns of Figure 8.12 and the barn-roof of Figure 9.1 (b) are two examples. Both require a novel construction line algorithm and a local drawing sequence for the volume composing parts that may occlude one another. The theory and practice of my painter's algorithm have been conceived such that local visibility in other simple volumes is easily handled, as demonstrated by the figures of this section which are rendered with my painter's algorithm.

The shapes made of non-planar surfaces are approximations of constructive solid geometry (CSG) primitives which have simple enough 2D projections, especially in central perspective: tangential relations of curvature with respect to the *pvp* and *dp* play a role in the local visibility of the surface. Thus, it is possible that projecting other operations, such as union, intersection, subtraction, into 2D, set operations commuting with projection would greatly diversify the simulated geometric shapes in my framework.

Adding new shapes requires extension of the shadow constructions. Figures 9.1 and 9.2 show that projected shadows can be produced by my existing framework in other simple circumstances, but additional work is needed to establish a general theory of projected shadows derived from 2D computations.

9.2.2 Multiple Perspectives

My framework has potential for investigating multiple perspectives in a picture as discussed in § 8.5. Pushing the research further would have applications to tasks where 2D images from different sources are combined. There have been good results in filling holes, image recolouring

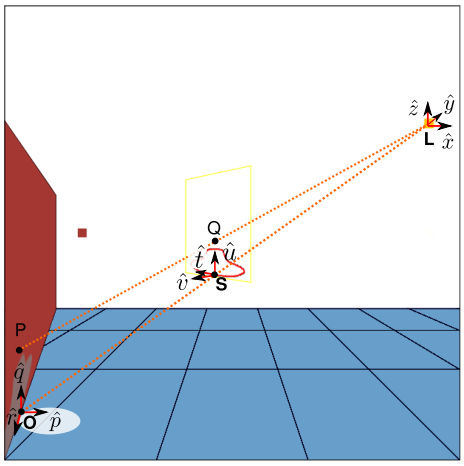
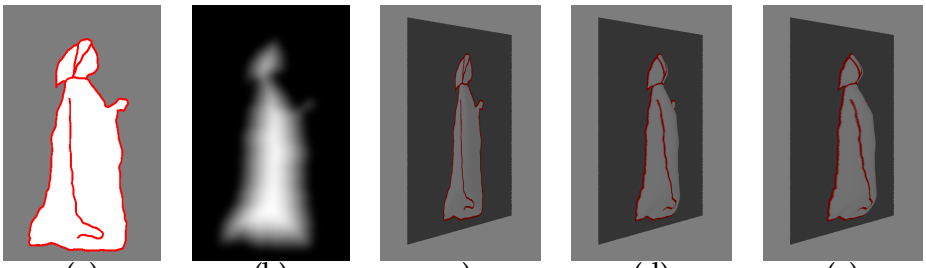


Figure 9.2: A general shadow theory of projected shadows computed in 2D is needed.

and even finding compatible images automatically, but little effort to alter pictures to improve composition. My description of De Chirico shows that the feeling of a composition depends strongly on the relative geometry of its parts. My framework shows it is possible to modify 2D geometry directly without needing 3D reconstruction, which will support tools for controlling image geometry in 2D.

A remaining challenge in understanding local perspective is to identify the visibility constraint: De Chirico bends space without breaking it. How much further could he go without the composition breaking into visual chaos? The picture is atypical. Compared to *The Ambassadors*, there is ample space between areas of conflicting geometry, but the discordance is not as subtle. My painter's algorithm has much flexibility. Can it be adapted to give consistent solutions on boundaries separating regions of inconsistent geometry? Or is such a goal irrelevant because users will inevitably demand creative control? And if so, what affordances are suitable for an artist user? Empirical investigation of the perceptual acceptability of inconsistent perspective, with subjects in the laboratory, not to mention the 'found experiments' [13] hanging on the walls of art galleries, is an important avenue for future work.

In effect, manipulating the *vp* locally produces volumes that do not align to the tiled floor. My painter's algorithm regardless processes each according to their initial position. Local *vp*s also compromise the occlusion produced by my painter's algorithm. It is possible to increase the resolution and rely on graphics hardware to make it work, but it would be better to find a more elegant solution. The framework user through the interactive feedback realizes when the scene visibility is compromised; whether the framework can predict for a specific scene when a manipulation is unsolvable is an interesting question.



(a) (b) c) (d) (e)

Figure 9.3: (a) The panel content representing a woman. (b) The heightmap texture for (a): whiter is more distant than darker; black has a local depth of zero; white is a local displacement, ϵ. (c) The rotated panel of the woman looks distorted since its rendering lacks local occlusion. (d) The same rotated panel where the heightmap is used to define and render the local relief of the interior surface represented on the panel. (e) same as (d) using a greater value for the maximum displacement of the heightmap.

9.2.3 Surface Details

Relief details would enhance the power of panels. Renaissance cartoons were purpose-created for specific locations in specific works and required only minimal adjustment. Modern computer graphics is more likely to incorporate found images which inevitably need more editing. Having rejected large panel rotations experimentally [40], I considered putting relief on panels, and discovered that a little hand-painted relief provides good results for disguising the flatness of images.

Silhouettes omit the geometric complexity of the interior of foreground objects, thereby falling short of cartoons. Artists can paint local reliefs on foreground objects with colouring and shading, giving them volumetric substance. These details painted simulate local occlusion, partial visibility bringing out simulated volume to planar representation. Self-occlusion is easy for artists, who without ado paint the first intersection, which is the one they see. I created a rendering algorithm for displacement mapping that works like an artist [41]. It steps along a visibility curve calculating the occlusions of a displaced surface by decomposing a 2D problem into a collection of 1D problems, which could be applied to panels.

Figure 9.3 illustrates the applicability of my displacement map rendering for improving my framework. Figure 9.3 (a) shows a panel with planar content. A heightmap (Figure 9.3 (b)) is painted freehand using 2D software to define the roughly cylindrical shape of the woman. Without relief rotating the panel too far produces a problem (Figure 9.3 (c)): the woman looks distorted because her flatness is obvious. The 2D displacement algorithm alleviates the problem by rendering the local occlusion (Figures 9.3 (d) & (e)). It is surprising how much a simply

drawn relief map improves the perceived geometry of the woman. It is especially interesting as the heightmap does not add many details for which the viewer is searching, but instead removes an anomaly that might otherwise distract the viewer.

Alternatively, panel content could be a local representation of a 3D model, which has been made beside my framework. It would be interesting to investigate adjusting 3D objects through the panel representation, where the panel for the 3D object is rerendered in local perspective in an off-screen buffer. This would allow framework users to modify the 3D object orientation. Manipulating 3D objects into the 2D context of the framework combines the advantages of realistically represented foreground volumes and of composition ease.

9.2.4 Other Directions

Another direction that particularly follow this thesis signature would be the application of my approach to realism for set design. In particular, stages are shallow, and set designers often wish to make them look deep, to differentiate the mood between the well-lit front of the stage and the dark mysterious back, for example. One standard methods of doing so is to build a ' half-perspective' design: the lateral walls converge and diminish in height, exaggerating foreshortening. The effect is often reinforced by a floor pattern, following a tiled floor converging to a vanishing point correct for the 'golden seat', where the designer and director sit during rehearsals. The result is a deep space, with many differentiated areas. The strong perspective overrules the absence of foreshortening of the actors. The ability to create in one volume the strong feeling of a different volume has many potential applications, from interior decorating to virtual reality.

This thesis demonstrates that perspective pictures can be created without leaving the canvas surface. The canvas plane is not limiting. With only few points realistic spatial coherence can be constructed: the pattern of depth is controlled by a geometric sequence revealed from intersections using pvp and dp lines. Small integer ratios derived from a single unit of measurement creates relationships in the 2D composition and in the 3D simulation. A common unit—the braccio, literally 'arm', which is equivalent to about 60 cm—underlies the visual rhythm of Renaissance paintings. The braccio, which originated in the cloth trade [37, p. 52], is identified as one-third of the average human height: Alberti specified the tiled floor construction not in term of the tile size but the horizon height, three braccia. It would be quite interesting to discover how perspective can embed the geometry of a 3D world within another using just few critical points on the canvas. Piero seems to have made such collage to diversify the shapes and spaces of his 2D composition within an out-normed 3D arrangement, as was discussed

for the *The Flagellation* (Figure 8.6). Given today's difficulty of integrating models that differ in resolution a recursive approach on the plane could be desirable to simplify and make effective 3D pictures with no limit. It is remarkable that, in the Renaissance geometry, was apprehended in a scalable approach obtaining resolution independence with small integers ratios from an humanly reference.

9.2.5 Closing Remark

This thesis, which limited itself to geometry, discovered an artists' view of geometry that has much to offer computer graphics, especially the lesson than nobody wherever they are in picture-making, can afford to lose sight of the ultimate goal, a picture that enriches the life of a viewer who is willing to contemplate it on its own terms. But a thesis on perspective geometry only scratches the surface. The challenge is two-fold: first, to understand the practices of artists and assimilate them to computer graphics, and second, to make them part of the education of every student in computer graphics. It is not enough to know algorithms, we must also learn to see, appreciate and understand.

Appendix A

Formalism Continued

Chapter 6 explained the Renaissance canvas space, $\mathbb{RCS}$, in which the 2D projective mappings function, and derived the matrices to transform general points in the perspective picture. This appendix provides more matrices that complete the manipulations of Chapter 5. Specifically, the following mappings are derived.

1. Between a point on a panel and the corresponding point at a different depth, which is needed when a panel is pushed in depth.
2. Between a point on an unrotated panel and the same point when the panel has undergone a rotation.
3. Between a shadowing point and its shadowed point on the ground plane.
4. Between a shadowing point and its shadowed point on an horizontal slice.
5. Between a shadowing point and its shadowed point intercepted on a vertical plane, which is parallel to the image plane.

A.1 General Point Bis

First, an alternate derivation for the general point mapping (§ 6.3) is presented. It shows the versatility by which the $\mathbb{RCS}$ can be considered. Figure A.1 illustrates the construction; the derivation follows. This equivalent construction is less intuitive as it uses a *downhill horizon*[1], but leads to a simpler mathematical derivation.

[1]The term comes from artistic perspective practice [17, Chapter VI].

The derivation is interesting because the ℝℂ𝕊 is considered differently: rather than lifting the baseline, the perspective ℝℂ𝕊 horizon is lowered to determine the perspective picture foreshortening. Physical ℝℂ𝕊 slices are not used, a different perspective ground whose horizon is pulled downward enables us to deduce the perspective ℝℂ𝕊 slice adhering to the original horizon.

A.1.1 Description

Figure A.1 shows the construction of the elevated tile corners on the slice $\mathbf{c}'$. The construction is in two parts.

1. Using a horizon lowered by $\mathbf{c}'$ from the original horizon, the tiled floor corners are constructed: the pvp and dp lower with the fixed baseline create a downhill tiled floor.
2. The downhill corners are raised by $\mathbf{c}'$ to be on perspective ℝℂ𝕊 slice $\mathbf{c}'$.

A.1.2 Derivation

Since the horizon is lowered by $\mathbf{c}'$, the mapping ϖ (Equation 6.1), which deduces a perspective ℝℂ𝕊 point from its physical ℝℂ𝕊 point, is replaced by $\varpi_{\mathbf{c}'}$. The matrix $\varpi_{\mathbf{c}'}$ is ϖ with the exception that the horizon height $\mathbf{E}_h$ is in this case $\mathbf{E}_h - \mathbf{c}'$.[2]

$$\varpi_{\mathbf{c}'} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h - \mathbf{c}' & 0 \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix}$$

The matrix $\pi_{\mathbf{c}'}$ that maps a physical ℝℂ𝕊 attachment point to the perspective ℝℂ𝕊 general point at physical height $\mathbf{c}'$ above its attachment point is given by the matrix multiplication $\tau_{\mathbf{c}'}\varpi_{\mathbf{c}'}$, where $\tau_{\mathbf{c}'}$ is a translation matrix that moves a point vertically by $\mathbf{c}'$. Thus,

$$\pi_{\mathbf{c}'} \sim \tau_{\mathbf{c}'}\varpi_{\mathbf{c}'} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \mathbf{c}' \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h - \mathbf{c}' & 0 \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix},$$

multiplies to the same transformation as $\pi_{\mathbf{c}'}$ (§ 6.3),

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h & -\mathbf{c}'\mathbf{E}_d \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix}. \tag{A.1}$$

[2] $\mathbf{c}'$ is both the physical height and the amount the horizon must be pulled down.

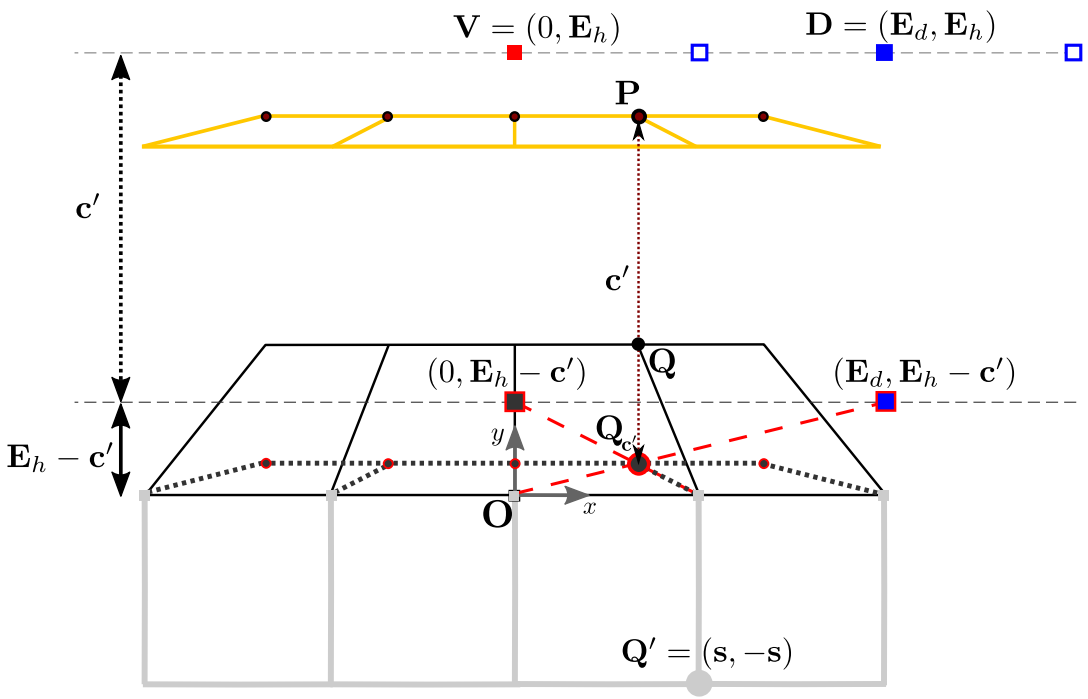


Figure A.1: The general point evaluation, $\mathbf{P}$, is done by constructing a different perspective $\mathbb{RCS}$ tiled floor (dotted dark grey) using a downhill horizon: the horizon lowered by $\mathbf{c}'$. Then this tiled floor result is raised by $\mathbf{c}'$ to produce the $\mathbb{RCS}$ perspective slice.

In effect, summarizing the derivations based on the two distinct geometric constructions gives the following relation:

$$\phi_{\mathbf{c}'}\tau_{\mathbf{c}'} = \tau_{\mathbf{c}'}\varpi_{\mathbf{c}'}\tau_{\mathbf{c}'}^{-1}\tau_{\mathbf{c}'} = \tau_{\mathbf{c}'}\varpi_{\mathbf{c}'} = \pi_{\mathbf{c}'}.$$

Interestingly this derivation is easier than the one of Chapter 6 (§ 6.3.2). The construction easier to use in practice is less congenial to derive mathematically. The inconvenience in practicality is that a lower tiled floor is estimated through moving the horizon. The transformations $\pi_{\mathbf{c}'}$ and $\Pi_{\mathbf{c}'}$ though, work for an arbitrary attachment location, and even a negative $\mathbf{c}'$.

A.2 Mappings of Lines

Mappings of lines permit checking matrices that transform the point, $\pi_{\mathbf{c}'}$ (Equation 6.3) among them. Such mappings are also useful in the derivation of complex $\mathbb{RCS}$ transformations: they are, for example, used in § A.3.3.

Duality is an important concept of projective geometry in its theory and its practice. For example, in planar projective geometry, the two following propositions are dual results of each other [85, p. 39–40].

Two distinct points lie on one and only one line.

Two distinct lines meet in one and only on point.

So far points have been used to derive the transformations. The relation between points and lines in homogeneous coordinates is useful: a point $\mathbf{Q}$ is on the line, $\mathscr{L} : (a,\ b,\ c)$, if and only if $\mathscr{L} \cdot \mathbf{Q} = 0$.

There is another practical advantage of planar projective geometry duality: the derived transformations have a double function as they also map lines. Precisely, the theorem is:

If the matrix M transforms the point P into P' then if $\mathscr{L}$ is a line then $\mathscr{L}M^{-1}$ is the equivalent transformed line, $\mathscr{L}'$ [23, p. 11].

Consider the transformation $\pi_{\mathbf{c}'}$ (Equation A.1), its inverse transforms a perspective $\mathbb{RCS}$ general point $\mathbf{P}$ at height $\mathbf{c}'$ to its attachment point, $\mathbf{Q}'$, on the physical $\mathbb{RCS}$ floor. In effect,

$$\begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix} \sim \pi_{\mathbf{c}'}^{-1} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_h + \mathbf{c}' & 0 & 0 \\ 0 & \mathbf{E}_d & -\mathbf{c}'\mathbf{E}_d \\ 0 & 1 & -\mathbf{E}_h \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix}. \tag{A.2}$$

But $\pi_{\mathbf{c}'}^{-1}$ also transforms any physical $\mathbb{RCS}$ floor line to a perspective $\mathbb{RCS}$ slice, at $\mathbf{c}'$. The simple vertical line of the physical $\mathbb{RCS}$ floor, $x = s$, whose homogeneous coordinates are $(1,\ 0,\ -s)$, is directly evaluated in the perspective $\mathbb{RCS}$ onto the $\mathbf{c}'$ slice.

$$\begin{pmatrix} 1 & 0 & -s \end{pmatrix} \pi_{\mathbf{c}'}^{-1} \sim \begin{pmatrix} \mathbf{E}_h - \mathbf{c}' & s & -s\mathbf{E}_h \end{pmatrix}.$$

As expected both the *pvp*, which is $\mathbf{V}$ here, and $(s, \mathbf{c}', 1)^T$ are points on this resulting line. Also, the perspective $\mathbb{RCS}$ horizon, $(0,\ 1,\ \mathbf{E}_h)$ is transformed by $\pi_{\mathbf{c}'}$ for any slice to the physical $\mathbb{RCS}$ line $(0,\ 0,\ k)$, the homogeneous row triplet of the line passing through all the points at infinity [23, p. 11].

Mappings of lines are natural checks, plus they are useful in derivations to consider points together, such as invariant or transformed axis.

A.3 Panel Manipulations

When the user of the framework manipulates a panel, its appearance changes as its properties are updated. Each panel has an attachment point and a local coordinate system, which combine to define its placement in the perspective picture: in the perspective $\mathbb{RCS}$, the attachment point is $\mathbf{Q}$, and the origin of the coordinate system $\mathbf{P}$; in the physical $\mathbb{RCS}$ the equivalent points are $\mathbf{Q}'$ and $\mathbf{P}'$. There is in addition an uniform scale, which defines the location of the panel's boundary corners. This section presents the formalism computations that encode the line intersection construction of § 5.4.2. This section describes how the formalism provides the panel affordances of moving a panel in the frontal plane and in depth, as well as rotating it about its vertical axis to change its orientation.

A.3.1 Frontal Plane Motion

Figure 5.10 (a) shows one way of moving a panel: motion in the frontal plane. The direct manipulation interface needs to redraw the panel as it is dragged. In the perspective $\mathbb{RCS}$, the principal change is the vertical offset separating the panel content from its attachment point: $\mathbf{P}_y$ moves leaving $\mathbf{Q}_y$ fixed. It is combined with a horizontal motion that translates $\mathbf{P}_x$ and $\mathbf{Q}_x$ together.

Frontal plane motion is an affordance that modifies the physical distance separating the panel from the ground plane[3]. When a point above the ground plane, $\mathbf{P}$, or its attachment point, $\mathbf{Q}$, are changed, the formalism must calculate the new physical height, $\mathbf{c}'$. (This reproduces the construction line evaluation in Figure 5.13.) Equation A.1 leads to the following equation for $\mathbf{c}'$ given $\mathbf{P}_y$ and $\mathbf{Q}'_y$:

$$\mathbf{c}' = \frac{\mathbf{Q}'_y(\mathbf{E}_h - \mathbf{P}_y) + \mathbf{E}_d\mathbf{P}_y}{\mathbf{E}_d}.$$

More usefully, the physical height $\mathbf{c}'$ can be retrieved from the perspective $\mathbb{RCS}$ points. Using Equation 6.4,

$$\mathbf{c}' = \frac{(\mathbf{P}_y - \mathbf{Q}_y)\mathbf{E}_h}{\mathbf{E}_h - \mathbf{Q}_y}.$$

Similar to other instances discussed in § 6.4, we observe that this direct computation in the perspective $\mathbb{RCS}$ produces a result independent of $\mathbf{E}_d$. Using perspective $\mathbb{RCS}$ points, there is

[3] A scaling in the image plane, Figure 5.11 (a), similarly changes it. It also changes the panel's height, the location of its boundary top edge.

no direct effect of $\mathbf{E}_d$ in defining the point $\mathbf{P}$. $\mathbf{E}_d$ has an effect only the location of $\mathbf{Q}$; for $\mathbf{P}$ the distance between $\mathbf{Q}$ and the horizon is sufficient.

A.3.2 Depth Motion

Figure 5.10 (b) shows how a panel moves in depth. The simulation translates the panel attachment point, $\mathbf{Q}$, parallel to the principal line of sight. Thus, $\mathbf{Q}$ is moved along the floor orthogonals direction. Doing so for an object fixed in physical height requires modification of the perspective $\mathbb{RCS}$ heights.

On the perspective $\mathbb{RCS}$, the attachment point moves from $\mathbf{Q}^1$ along its line to pvp to $\mathbf{Q}^2$, a distance δ_y in the y-direction. Without loss of generality δ_y may be expressed as a multiple of the vertical distance separating $\mathbf{Q}^1$ from the horizon: $\delta_y = k(\mathbf{E}_h - \mathbf{Q}^1_y)$. Because the resulting $\mathbf{Q}^2$ lies on the line $\mathbf{VQ}^1 : (\mathbf{Q}^1_y - \mathbf{E}_h, \quad -\mathbf{Q}^1_x, \quad \mathbf{Q}^1_x\mathbf{E}_h)$ and the y-coordinate of $\mathbf{Q}^2$ is $\mathbf{Q}^1_y - k(\mathbf{E}_h - \mathbf{Q}^1_y)$, we have the following transformation that moves a point in depth on the perspective $\mathbb{RCS}$:

$$\mathbf{Q}^2 \sim \begin{pmatrix} 1-k & 0 & 0 \\ 0 & 1-k & k\,\mathbf{E}_h \\ 0 & 0 & 1 \end{pmatrix} \mathbf{Q}^1.$$

Then, the panel foreshortened appearance is adjusted. For example, $|\mathbf{P}_1 - \mathbf{Q}_1|$ corresponds to a physical value $\mathbf{c}'$ that stays fixed. Using $\Pi_{\mathbf{c}'}$ (Equation 6.4) an attachment point is mapped to the point above it at height $\mathbf{c}'$, thus $\mathbf{P}_2$ is directly derived by the following transformation, $\Gamma_{\mathbf{c}'}$:

$$\mathbf{P}^2 \sim \Gamma_{\mathbf{c}'}\mathbf{Q}^1. \sim \begin{pmatrix} (1-k)\mathbf{E}_h & 0 & 0 \\ 0 & (1-k)(\mathbf{E}_h - \mathbf{c}') & (k(\mathbf{E}_h - \mathbf{c}') + \mathbf{c}')\mathbf{E}_h \\ 0 & 0 & \mathbf{E}_h \end{pmatrix} \mathbf{Q}^1.$$

Γ is the general transformation calculating a perspective $\mathbb{RCS}$ point on or above the floor after motion in depth, $\delta_y = k(\mathbf{E}_h - \mathbf{Q}^1_y)$, in terms of its starting perspective $\mathbb{RCS}$ attachment point. It is essential to update the panel origin position and the transformation can also be used to find the new location of the panel's top edge[4]. The perspective $\mathbb{RCS}$ transformation, Γ, is again independent of $\mathbf{E}_d$.

[4] Alternatively, the ratio of $|\mathbf{P}^1 - \mathbf{Q}^1|$ to $|\mathbf{P}^2 - \mathbf{Q}^2|$ can be used to update the scaling factor applied to the panel content: its boundary changes with the push in depth.

A.3.3 Rotation

The formalism must have a transformation that changes the panel appearance when a user rotates a panel around its central vertical axis. My construction for this affordance needs to be checked: illustrated in Figure 5.16, the panel boundary into which the panel content is warped.

Derivation. To derive the transformation that follows the construction line computations for rotating a panel (Figures 5.15 and 5.16) two geometric properties suffice.

1. The central vertical axis of the panel stays fixed.
2. The horizontal line through the panel corner attachment points is rotated by angle θ.

The first property translates to a trivial relation from the previous derivations. The second property creates new relations. To locate in the perspective $\mathbb{RCS}$ the floor points, $\mathbf{L}_\theta$ and $\mathbf{R}_\theta$ in Figure 5.15 (b), their physical $\mathbb{RCS}$ counterparts are used. In effect, $\mathbf{L}'_\theta$ and $\mathbf{R}'_\theta$ are on the line, $\mathscr{L}'_\theta$ on the physical floor, which is rotated by $-\theta$ around $\mathbf{Q}'$. $\mathscr{L}'_\theta$ has the following homogeneous coordinates:

$$\mathscr{L}'_\theta : \Big((\mathbf{E}_h - \mathbf{Q}_y) \tan\theta \qquad \mathbf{E}_h - \mathbf{Q}_y \qquad -\mathbf{Q}_x \mathbf{E}_h \tan\theta + \mathbf{Q}_y \mathbf{E}_d \Big) .$$

The points on it can be expressed in term of the physical distance $\mathbf{l}'$ separating them from the center of rotation $\mathbf{Q}'$. In effect, the points $\mathbf{Q}'_{\mathbf{l}',\theta}$ are expressed by

$$\mathbf{Q}'_{\mathbf{l}',\theta} \sim \begin{pmatrix} -\mathbf{l}' \, \mathbf{E}_h \cos\theta - \mathbf{Q}_x \mathbf{E}_h \\ \mathbf{l}' \, \mathbf{E}_h \sin\theta + \mathbf{Q}_y \mathbf{E}_d \\ \mathbf{Q}_y - \mathbf{E}_h \end{pmatrix} .$$

These physical $\mathbb{RCS}$ floor points, $\mathbf{Q}'_{\mathbf{l}',\theta}$, transform to perspective $\mathbb{RCS}$ floor points by applying ϖ (Equation 6.1). They are

$$\mathbf{Q}_{\mathbf{l}',\theta} \sim \begin{pmatrix} \mathbf{l}' \, \mathbf{E}_d \cos\theta + \mathbf{Q}_x \mathbf{E}_d \\ \mathbf{l}' \, \mathbf{E}_h \sin\theta + \mathbf{Q}_y \mathbf{E}_d \\ \mathbf{l}' \, \sin\theta + \mathbf{E}_d \end{pmatrix} .$$

These definitions are sufficient to derive the matrix that performs rotation in the perspective $\mathbb{RCS}$. The two mappings below, expressing correspondence between lines, determine the rotation matrix. The first mapping keeps the points on the vertical axis through $\mathbf{Q}$ fixed: Equation A.1 is used to express all the points $\mathbf{P}$ in function of $\mathbf{Q}$ for all $\mathbf{c}'$. The second one moves the floor points of the panel base projection: initially aligned horizontally to the oblique alignment they assume after the panel rotation.

$$\begin{pmatrix} \mathbf{Q}_x\mathbf{E}_h \\ \mathbf{E}_h(\mathbf{Q}_y+\mathbf{c}')-\mathbf{c}'\mathbf{Q}_y \\ \mathbf{E}_h \end{pmatrix} \mapsto \alpha \begin{pmatrix} \mathbf{Q}_x\mathbf{E}_h \\ \mathbf{E}_h(\mathbf{Q}_y+\mathbf{c}')-\mathbf{c}'\mathbf{Q}_y \\ \mathbf{E}_h \end{pmatrix}, \forall \mathbf{c}' \quad \text{and}$$

$$\begin{pmatrix} \mathbf{Q}_x+\mathbf{l}' \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \mapsto \beta \begin{pmatrix} \mathbf{l}'\,\mathbf{E}_d\cos\theta+\mathbf{Q}_x\mathbf{E}_d \\ \mathbf{l}'\,\mathbf{E}_h\sin\theta+\mathbf{Q}_y\mathbf{E}_d \\ \mathbf{l}'\,\sin\theta+\mathbf{E}_d \end{pmatrix}, \forall \mathbf{l}'.$$

Using them in a derivation similar to the one described in Section 6.2 determines the matrix Θ below. Θ transforms the panel points projected on the floor, i.e., their attachment definitions, each known in terms of its lateral displacement to the panel attachment point $\mathbf{Q}$, to their new floor locations after the panel has been rotated by θ around the vertical axis through $\mathbf{Q}$:

$$\mathbf{Q}_{\mathbf{l}',\theta} \sim \Theta\ \mathbf{Q}_{\mathbf{l}'} \sim \begin{pmatrix} \mathbf{E}_d\cos\theta & 0 & \mathbf{Q}_x\mathbf{E}_d(1-\cos\theta) \\ \mathbf{E}_h\sin\theta & \mathbf{E}_d & -\mathbf{Q}_x\mathbf{E}_h\sin\theta \\ \sin\theta & 0 & \mathbf{E}_d-\mathbf{Q}_x\sin\theta \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x+\mathbf{l}' \\ \mathbf{Q}_y \\ 1 \end{pmatrix}.$$

Θ evaluates $\mathbf{L}_\theta$ and $\mathbf{R}_\theta$ of Figure 5.16: the attachment points for the vertical edges of the panel after the rotation, known initially by their lateral displacement from $\mathbf{Q}$, $\pm\mathbf{l}' = \pm\frac{w}{2}$, where w is the physical panel width (possibly scaled). But this matrix Θ extends to a general version, $\Theta_{\mathbf{c}'}$, which applies to perspective $\mathbb{RCS}$ general points $\mathbf{P}$: the points $\mathbf{c}'$ above the floor are included using $\Pi_{\mathbf{c}'}$ (Equation 6.4).

$$\mathbf{P}_{\mathbf{l}',\theta} \sim \Pi_{\mathbf{c}'}\,\Theta\ \mathbf{Q}_{\mathbf{l}'} \sim \Theta_{\mathbf{c}'}\ \mathbf{Q}_{\mathbf{l}'}$$

$$\mathbf{P}_{\mathbf{l}',\theta} \sim \begin{pmatrix} \mathbf{E}_d\mathbf{E}_h\cos\theta & 0 & \mathbf{Q}_x\mathbf{E}_d\mathbf{E}_h(1-\cos\theta) \\ \mathbf{E}_h^2\sin\theta & \mathbf{E}_d(\mathbf{E}_h-\mathbf{c}') & \mathbf{E}_h(\mathbf{c}'\mathbf{E}_d-\mathbf{Q}_x\mathbf{E}_h\sin\theta) \\ \mathbf{E}_h\sin\theta & 0 & \mathbf{E}_h(\mathbf{E}_d-\mathbf{Q}_x\sin\theta) \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x+\mathbf{l}' \\ \mathbf{Q}_y \\ 1 \end{pmatrix}.$$

Using $\Theta_{\mathbf{c}'}$ with two distinct values of $\mathbf{c}'$ permits us to deduce the four corners of the rotated panel boundary, the bottom edge and top one being elevated by different physical height. Thus, the transformation $\Theta_{\mathbf{c}'}$ replicates the construction line evaluations that rotate the panel boundary.

Comments. In $\Theta_{\mathbf{c}'}$ a global $\mathbf{E}_h$ factor exists: if $\mathbf{c}'$ had been expressed has a ratio of $\mathbf{E}_h$ (like to δ_y in § A.3.2), the factor would have been apparent and cancelling it would produce a simpler matrix. Regardless, for the panel rotated appearance the resulting matrix has clearly a dependence on the principal perspective: it depends on both $\mathbf{E}_h$ and $\mathbf{E}_d$.

When a panel is rotated by θ ($\neq n\pi, n \in \mathbb{N}$), its non-vertical edges converge to the horizon point, vp_θ, as indicated in Figure 5.16. Panels rotated by the same angle θ share that point of convergence as illustrated in Figure 5.12. Indeed there is an infinite point $(-1,\ 0,\ 0)^T$ contained in all panel planes that are parallel to the image plane. For each rotation angle θ the infinite point is transformed by $\Theta_{\mathbf{c}'}$ to $(\mathbf{E}_d / \tan\theta,\ \mathbf{E}_h,\ 1)^T$: the location vp_θ on the horizon which is independent of the panel attachment point $\mathbf{Q}$, but depends on the distance of the eye to the image plane $\mathbf{E}_d$ (as was discussed in the end of Section 5.4.2). vp_θ characterizes panels rotated by a same angle regardless of their positions[5].

Finally, computing algebraically brings the usual robustness advantage. Compared to the difficulty occurring in the construction line evaluation for a 90° rotation when the panel's origin is aligned with the pvp, the algebraic computation locates all the panel's boundary corners, which are aligned on the image central axis (even though the appearance is not actually desired). This is a good example of a feature that must be addressed in a computer implementation but which is unnecessary for a human who will catch it (at run-time) and respond easily case-by-case.

A.4 Shadow Mappings

In this section, the projective transformations replicating the geometric constructions for shadows described in Section A.4 are derived. These matrices calculate in the perspective $\mathbb{RPCS}$ the projected shadows of points given the projected position of the light on the canvas, that is its perspective $\mathbb{RPCS}$ vanishing point and the corresponding light source.

First the basic projection of a perspective $\mathbb{RPCS}$ general point to its shadow on the ground plane is given. This mapping creates shadows of panel content on the ground plane. Then projections to other planes are developed: shadows cast onto planes parallel and above the ground plane are simple generalizations; shadows cast on vertical planes parallel to the image

[5]The formalism permits alternate derivations. vp_θ is similarly derived using transformed lines (in the style of Section A.2): the physical line, $\mathscr{L}'_\theta$, transforms to its perspective $\mathbb{RPCS}$ version $\mathscr{L}_\theta$; this latter intersection with the horizon, $(0,\quad 1,\quad \mathbf{E}_h)$ also determines vp_θ.

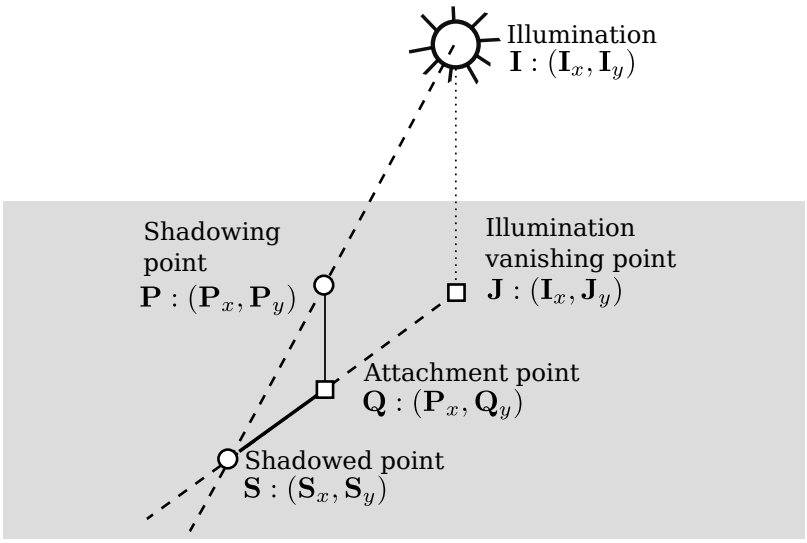


Figure A.2: A shadowing point creating a shadowed point on the ground plane. The illumination vanishing point and the attachment point are needed for the construction. The following labels are used in the derivation: the light source, **I**, its vanishing point on the ground, **J**, the shadowing point above the ground **P**, with its attachment point **Q**, and the resulting shadowed point, **S**, on the ground.

plane show the power of the formalism. The basic construction that depicts 2D ground shadows is proven to be optically correct in Appendix B.

The notation is extended to include the illumination-related points (Figure A.2). On the perspective $\mathbb{RCS}$ the light source is labeled **I** and its vanishing point on the ground plane **J**. Applied to a shadowing point above the ground plane **P**, with its attachment point, **Q**, the light creates its shadowed point, **S**, on the ground plane.

A.4.1 On the Ground Plane

The shadow effects privileged in my framework are those incurred by illuminations that are explicitly positioned within the canvas area, as artists often emphasize those: nearby point lights such as a light bulb on a lamp post, or distant lights low in the sky, such as the sun at down or dusk create the most evocative shadow effects.

In evaluating projected shadows the transformation that maps a point $\mathbf{P}$, which has attachment point $\mathbf{Q}$, to $\mathbf{S}$, where its shadow falls is first needed. The shadow lies at the intersection of the ground ray passing by $\mathbf{Q}$ and the light ray passing by $\mathbf{P}$. As the ground ray, $\mathscr{L}_g$, has the following homogeneous coordinates,

$$\mathscr{L}_g : \quad \begin{pmatrix} \mathbf{Q}_y - \mathbf{J}_y & \mathbf{I}_x - \mathbf{P}_x & \mathbf{P}_x\mathbf{J}_y - \mathbf{Q}_y\mathbf{I}_x \end{pmatrix}, \tag{A.3}$$

and the light ray, $\mathscr{L}_l$, has the following homogeneous coordinates,

$$\mathscr{L}_l : \quad \begin{pmatrix} \mathbf{P}_y - \mathbf{I}_y & \mathbf{I}_x - \mathbf{P}_x & \mathbf{P}_x\mathbf{I}_y - \mathbf{P}_y\mathbf{I}_x \end{pmatrix}, \tag{A.4}$$

and we have $\mathscr{L}_g \cdot \mathbf{S} = \mathscr{L}_l \cdot \mathbf{S} = 0$, when $\mathbf{I}_x \neq \mathbf{P}_x$[6], the shadowed point is

$$\mathbf{S} \sim \begin{pmatrix} \mathbf{P}_x(\mathbf{I}_y - \mathbf{J}_y) - \mathbf{P}_y\mathbf{I}_x + \mathbf{Q}_y\mathbf{I}_x \\ -\mathbf{P}_y\mathbf{J}_y + \mathbf{Q}_y\mathbf{I}_y \\ -\mathbf{P}_y + \mathbf{I}_y - \mathbf{J}_y + \mathbf{Q}_y \end{pmatrix}.$$

Thus the matrix, Ψ, applied to $\mathbf{P}$, calculates the location of its ground shadow $\mathbf{S}$:

$$\begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \\ 1 \end{pmatrix} \sim \Psi \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{I}_y - \mathbf{J}_y & -\mathbf{I}_x & \mathbf{Q}_y\mathbf{I}_x \\ 0 & -\mathbf{J}_y & \mathbf{Q}_y\mathbf{I}_y \\ 0 & -1 & \mathbf{I}_y - \mathbf{J}_y + \mathbf{Q}_y \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix}. \tag{A.5}$$

The light position points ($\mathbf{I}$ and $\mathbf{J}$) are parameters of the transformation, as is the coordinate $\mathbf{Q}_y$. The dependence of Ψ on $\mathbf{Q}_y$ and its independence of $\mathbf{Q}_x$ makes the expression convenient for evaluating the ground plane shadow of a panel, which requires the calculation of four corner shadow points.

Indeed, the four corners of a panel in its default orientation all have the same value of $\mathbf{Q}_y$, and are transformed by the same shadow matrix: Ψ does not vary from point to point, only from panel to panel. When a panel is not in its default orientation, two matrices are necessary to compute the projection of the corner points on either side of the panel central axis. Then from the mappings of the four points and their projections on the ground plane, there is a unique projective transformation mapping between the two quadrilaterals. This map projects the shadow mask to the ground plane, where it appears as an anamorphic projection of the opaque pixels of the panel content.

The mapping depends only on 2D points on the perspective $\mathbb{RCS}$; no 3D calculations are necessary. The 3D properties of shadow projection are calculated using the attachment/vanishing

[6]When $\mathbf{I}_x = \mathbf{P}_x$ the light and ground rays coincide making $\mathbf{S}$ impossible to determine.

points, which themselves are defined within the environment of perspective picture. The amazing part being that this shadow appearance is explicitly independent of the perspective $\mathbb{RCS}$ parameters, $\mathbf{E}_d, \mathbf{E}_h$ being absent from the mapping. The attachment (also vanishing) points encode just enough 3D information to make the calculation possible, and encode it in a way that makes it accessible in 2D. This observation is an important contribution of this thesis. The perspective is part of each computation and evaluation, comprised in each, rather than being an afterthought. As such, it easily accepts continuous input from an user, unlike a photograph, which projects to a 2D image only after objects have been placed in a 3D space.

The matrix Ψ cannot be used directly if the shadowing point is higher above the ground plane than is the light (Figure A.3). As described in the special cases of § 5.5.3 and illustrated in practice in Figure 5.20 (b), the intersection of the light and ground rays is then the projection to the canvas of a point behind the viewer in 3D (Figure A.3 (b)). To remove this wrap around, only parts of shadows that are behind the image plane should be considered (Figure A.3 (a)).

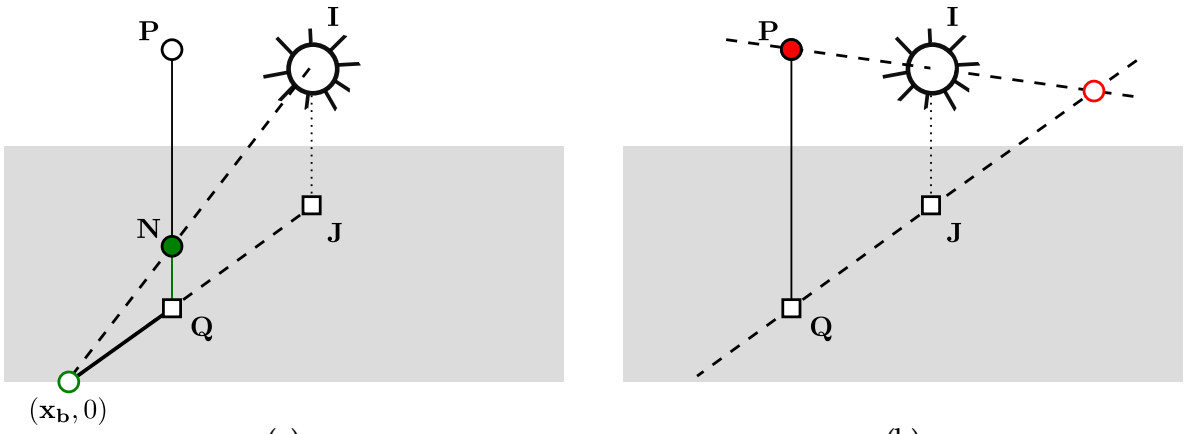


Figure A.3: Projected ground shadow geometry when the shadowing point is higher than the light. (a) the right effect is obtained by using only the part behind the image plane. (b) the geometry unclipped, the ground ray and light ray intersect on the wrong side of the light.

If $\mathbf{I}_y - \mathbf{J}_y < \mathbf{P}_y - \mathbf{Q}_y$, Ψ is not used directly on the point $\mathbf{P}$. To discover the computation, imagine the simple case of the shadow projection of the line segment $\mathbf{QP}$, for which we determine the necessary and sufficient clip to isolate the part of the shadow behind the image plane. Since the ground ray, $\mathscr{L}_g$, intersects the baseline at $(\mathbf{x}_b,\ 0,\ 1)^T$, we know that $\mathbf{x}_b = \frac{\mathbf{Q}_y\mathbf{I}_x - \mathbf{Q}_x\mathbf{J}_y}{\mathbf{Q}_y - \mathbf{J}_y}$, which is defined by the following matrix.

$$\begin{pmatrix} \mathbf{x}_b \\ 0 \\ 1 \end{pmatrix} \sim \xi \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{J}_y & -\mathbf{I}_x & 0 \\ 0 & 0 & 0 \\ 0 & -1 & \mathbf{J}_y \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix}. \tag{A.6}$$

Then the inverse of Ψ is used to determine the point on the axis **QP**, **N**, which corresponds to the intersection of the shadow with the baseline.

$$\Psi^{-1} \sim \begin{pmatrix} \mathbf{J}_y - \mathbf{Q}_y & -\mathbf{I}_x & \mathbf{I}_x\mathbf{Q}_y \\ 0 & -\mathbf{I}_y + \mathbf{J}_y - \mathbf{Q}_y & \mathbf{I}_y\mathbf{Q}_y \\ 0 & -1 & \mathbf{J}_y \end{pmatrix}. \tag{A.7}$$

Equations A.6 and A.7 reproduce the construction method of Figure 5.21. It is most effective to combine the two operations into a single matrix. Determining the general point **N** on the **PQ** axis that corresponds to the intersection with the baseline, based simply on the attachment point, **Q**, and the light position (**I**, and **J**) is handy. The following matrix does it exactly.

$$\begin{pmatrix} \mathbf{N}_x \\ \mathbf{N}_y \\ 1 \end{pmatrix} \sim \Psi^{-1}\xi \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{J}_y & 0 & 0 \\ 0 & \mathbf{I}_y & 0 \\ 0 & 0 & \mathbf{J}_y \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix}. \tag{A.8}$$

The matrix is simple and $\mathbf{N}_x = \mathbf{Q}_x$ because it must lie on the line **PQ**. The height to be clipped is a simple relationship between the y-coordinates of the points that define the illumination on the perspective $\mathbb{RCS}$. Thus, when the panel height is greater than to the light source height, only the region that shades behind the image plane is drawn, avoiding canvas areas affected by 3D points behind the viewer. In effect, the vertical image plane is the near plane, clipping the projected shadows on the ground plane. Using Equation A.8, the clipped region on a panel in its default orientation is determined early: $\mathbf{N}_y$ is compared to the height of the panel to determine which part of the panel casts the visible shadow. If the panel origin is high above the ground plane, the entire ground plane shadow may be clipped. For a rotated panel, the computations with respect to the baseline have to be done at least twice, when the panel base is on the ground plane once for each of the two top vertices of of the panel boundary. Finally, also keep in mind that a similar process is to be applied with respect to the horizon, when a pseudo light is low, for example.

The methods of this section generalize to calculate the shadows cast by the faces of blocks.

A.4.2 On Slices

The computations for shadows on the ground plane are now generalized to other surfaces. Figure 5.23 shows a construction line method for shadows that fall on the top face of a volume, an elevated plane parallel to the ground plane. The construction is similar to the one that produces ground plane shadows, except that finding the elevated attachment and light vanishing points on this shadow plane is first required. Then, it is only a matter to intersect the light rays with rays on the *elevated ground*. Ideally, the formalism integrates the elevated shadows within a computation that handles the ground plane as a special case. The perspective $\mathbb{RCS}$ horizontal slices (Chapter 6) are the planes artists consider as elevated ground planes.

Consider a point light casting a shadow on one of these slices, at physical height, $\mathbf{c}'$, above the ground plane; the vertical coordinates of its attachment point, $\mathbf{Q}_y$, and of the light vanishing point, $\mathbf{J}_y$, must to be elevated to the slice in the perspective $\mathbb{RCS}$: the elevated points are $\mathbf{Q}_y^\star$ and $\mathbf{J}_y^\star$, defined by Π (Equation 6.4). These values are substituted in Ψ (Equation A.5): $\mathbf{J}_y \to \mathbf{J}_y^\star$ and $\mathbf{Q}_y \to \mathbf{Q}_y^\star$. This new matrix, $\Psi_{\mathbf{c}'}^\star$,

$$\Psi_{\mathbf{c}'}^\star \sim \begin{pmatrix} \mathbf{I}_y - \mathbf{J}_y^\star & -\mathbf{I}_x & \mathbf{I}_x\mathbf{Q}_y^\star \\ 0 & -\mathbf{J}_y^\star & \mathbf{I}_y\mathbf{Q}_y^\star \\ 0 & -1 & \mathbf{I}_y - \mathbf{J}_y^\star + \mathbf{Q}_y^\star \end{pmatrix}, \tag{A.9}$$

calculates the projected shadow $\mathbf{S}^\star$ on the elevated plane.

It is possible to determine the general matrix as a function of the initial points $\mathbf{Q}$ and $\mathbf{J}$ on the ground plane, which evaluates any shadow on perspective $\mathbb{RCS}$ slices. In effect, $\mathbf{Q}_y^\star$ and $\mathbf{J}_y^\star$, can be computed using $\Pi_{\mathbf{c}'}$ (Equation 6.4), with

$$\mathbf{P} \sim \Pi_{\mathbf{c}'}\mathbf{Q} \sim \begin{pmatrix} \mathbf{Q}_x \\ \frac{\mathbf{E}_h\mathbf{Q}_y + \mathbf{c}'(\mathbf{E}_h - \mathbf{Q}_y)}{\mathbf{E}_h} \\ 1 \end{pmatrix}$$

giving

$$\mathbf{Q}_y^\star \sim \frac{\mathbf{E}_h\mathbf{Q}_y + \mathbf{c}'(\mathbf{E}_h - \mathbf{Q}_y)}{\mathbf{E}_h} \qquad \text{and} \qquad \mathbf{J}_y^\star \sim \frac{\mathbf{E}_h\mathbf{J}_y + \mathbf{c}'(\mathbf{E}_h - \mathbf{J}_y)}{\mathbf{E}_h}.$$

Replacing them in Equation A.9, we obtain the following general matrix, $\Psi_{\mathbf{c}'}$,

$$\Psi_{\mathbf{c}'} \sim \begin{pmatrix} \mathbf{E}_h(\mathbf{I}_y - \mathbf{J}_y) - \mathbf{c}'(\mathbf{E}_h - \mathbf{J}_y) & -\mathbf{E}_h\mathbf{I}_x & \mathbf{I}_x(\mathbf{Q}_y\mathbf{E}_h + \mathbf{c}'(\mathbf{E}_h - \mathbf{Q}_y)) \\ 0 & -\mathbf{E}_h\mathbf{J}_y - \mathbf{c}'(\mathbf{E}_h - \mathbf{J}_y) & \mathbf{I}_y(\mathbf{Q}_y\mathbf{E}_h + \mathbf{c}'(\mathbf{E}_h - \mathbf{Q}_y)) \\ 0 & -\mathbf{E}_h & \mathbf{E}_h(\mathbf{I}_y - \mathbf{J}_y + \mathbf{Q}_y) - \mathbf{c}'(\mathbf{Q}_y - \mathbf{J}_y) \end{pmatrix} \tag{A.10}$$

which projects the general shadowing point, $\mathbf{P}$, to its shadowed point, $\mathbf{S}_{\mathbf{c}'}$, on the $\mathbf{c}'$ slice, such that $\mathbf{S}_{\mathbf{c}'} \sim \Psi_{\mathbf{c}'}\mathbf{P}$. $\Psi_{\mathbf{c}'}$ is the general matrix that includes $\Psi_{\mathbf{c}'} \sim \Psi$ for the special case when $\mathbf{c}' = 0$. While shadows on the ground plane are independent of the scene perspective, $\mathbf{E}_d$ and $\mathbf{E}_h$, the height of the viewpoint, $\mathbf{E}_h$, must be taken into account when evaluating with $\mathbf{c}' \neq 0$ because shadows on arbitrary slices have appearances that are affected by the scene perspective. The appearance of blocks and particularly of their top faces depends on the height of the horizon, as do the appearances of shadows that fall on them.

As for Ψ, the inverse of $\Psi_{\mathbf{c}'}$ is required for clipping. Using the above substitutions, $\Psi_{\mathbf{c}'}^{-1}$ is similarly determined

$$\Psi_{\mathbf{c}'}^{-1} \sim \begin{pmatrix} (\mathbf{E}_h - \mathbf{c}')(\mathbf{J}_y - \mathbf{Q}_y) & -\mathbf{E}_h\mathbf{I}_x & \mathbf{I}_x(\mathbf{Q}_y\mathbf{E}_h + \mathbf{c}'(\mathbf{E}_h - \mathbf{Q}_y)) \\ 0 & \mathbf{E}_h(\mathbf{J}_y - \mathbf{Q}_y - \mathbf{I}_y) - \mathbf{c}'(\mathbf{J}_y - \mathbf{Q}_y) & \mathbf{I}_y(\mathbf{Q}_y\mathbf{E}_h + \mathbf{c}'(\mathbf{E}_h - \mathbf{Q}_y)) \\ 0 & -\mathbf{E}_h & \mathbf{E}_h\mathbf{J}_y + \mathbf{c}'(\mathbf{E}_h - \mathbf{J}_y) \end{pmatrix} \quad \text{(A.11)}$$

$\Psi_{\mathbf{c}'}^{-1}$ is used to generalize baseline clipping to any slice above the baseline. The matrix in Equation A.8 is re-parametrized to include $\mathbf{c}'$, extending the baseline clipping used for projections on the ground plane to any joint line clipping for their respective slice projections. $\mathbf{N}_{\mathbf{c}'}$ is now the point on the line $\mathbf{PQ}$ that corresponds to the intersection with the joint line.

$$\mathbf{N}_{\mathbf{c}'} \sim \begin{pmatrix} \mathbf{J}_y & 0 & 0 \\ 0 & \mathbf{I}_y - \mathbf{c}' & \mathbf{c}'\mathbf{J}_y \\ 0 & 0 & \mathbf{J}_y \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix}. \quad \text{(A.12)}$$

Matrix $\Psi_{\mathbf{c}'}$ and Equation A.12 are used to clip the shadows of panels for the top faces of blocks. They provide the projections corresponding to panels' regions that are visible on those perspective $\mathbb{RCS}$ slices. As described above the shadows sometimes must be clipped at the image plane or at the horizon.

While the formalism clips at the image plane and at the horizon plane to compute the shadow projection in the perspective $\mathbb{RCS}$, further clipping is applied to ensure projected shadows only appear to fall on opaque surfaces in the perspective picture. In effect, the rendering of the perspective picture processes projected shadows through clips of block faces. For example, a block top face drawn in the perspective picture is a finite area, while a projected shadow atop its surface is computed with only the consideration of near and far planes, the rendering uses block face clips to guarantee the projected shadow remains in contact with the visible surface.

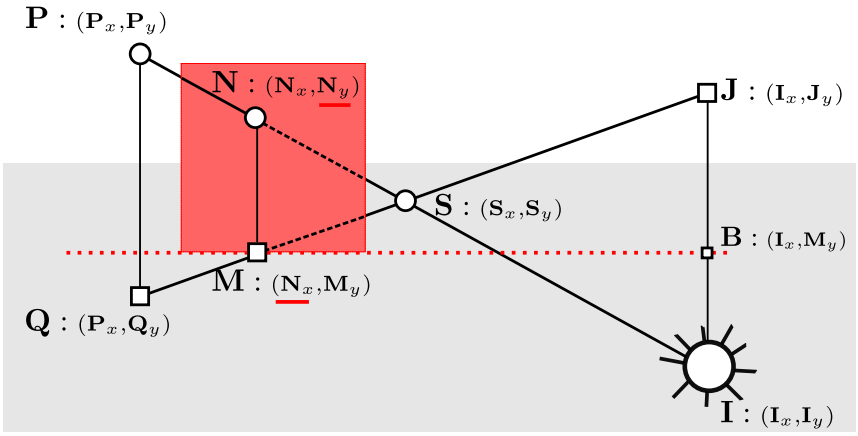


Figure A.4: A light ray from the pseudo-sun intersecting a wall (a bounded part of a vertical plane parallel to the image plane) at a depth $\mathbf{M}_y$. The formalism determines the shadowed point $\mathbf{N}$, which is the projection onto this wall of a general shadowing point $\mathbf{P}$ with respect to the light position, $(\mathbf{I}, \mathbf{J})$.

A.4.3 On Vertical Planes

Now, the capability of the formalism is further demonstrated by developing calculations for shadows on vertical surfaces parallel to the image plane. An example occurs in Figure 5.24, where the actual illumination, which appears as a pseudo-light on the perspective $\mathbb{R}$CS, is behind the viewer: the pseudo-sun is truly an anti-sun as it absorbs light that travels from illuminated objects to it. The shadow of a panel falls partly on the ground plane and continues onto a vertical wall deeper in the scene. To find a shadow on a vertical plane placed at a known scene depth, the projection of a general point in the panel content, $\mathbf{P}$, with attachment point $\mathbf{Q}$, onto the plane for the light location needs to be known.

Figure A.4 illustrates an instance of the geometry in play in such a configuration. The physical depth of the vertical plane is $\mathbf{M}_y$. The question is where does the light ray to the general point $\mathbf{P}$, which intersects with the ground plane at $\mathbf{S}$, intersect above the ground ray within the sub-plane of a vertical plane: the shadowed point $\mathbf{N}$.

The attachment point of the shadowed point $\mathbf{N}$ is on the ground plane at $\mathbf{M}$, with $\mathbf{M}_y$ known, thus, $\mathbf{N}_x = \mathbf{M}_x$ must be determined first. Figure A.5 shows the relevant geometry, emphasizing the similar triangles $\triangle\mathbf{JCQ}$ and $\triangle\mathbf{JMB}$, which permit to evaluate $\mathbf{M}_x$. Alternatively, passing through $\mathbf{Q}$ from $\mathbf{J}$, the ground ray $\mathscr{L}_g$ (Equation A.3) determines $\mathbf{M}_x$ as we have

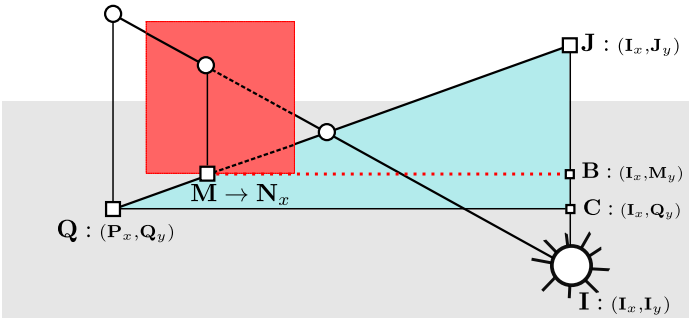


Figure A.5: Geometry for determining the lateral position, $\mathbf{N}_x$. $\mathbf{P}$ projects to $\mathbf{N}$ on a wall situated at perspective $\mathbb{RCS}$ depth $\mathbf{M}_y$, parallel to the image plane. The center of projection is the light position, $(\mathbf{I}, \mathbf{J})$. $\mathbf{N}_x$ is the intersection of the ground ray $\mathbf{QJ}$ with the line parallel to the baseline at depth $\mathbf{M}_y$, $\mathbf{MB}$.

$$\begin{aligned} \mathscr{L}_g \cdot \mathbf{M} &= 0 \\ (\mathbf{Q}_y - \mathbf{J}_y)\mathbf{M}_x + (\mathbf{I}_x - \mathbf{P}_x)\mathbf{M}_y + \mathbf{P}_x\mathbf{J}_y - \mathbf{Q}_y\mathbf{I}_x &= 0 \\ \mathbf{M}_x = \mathbf{N}_x &= \frac{\mathbf{P}_x(\mathbf{J}_y - \mathbf{M}_y) + \mathbf{I}_x(\mathbf{M}_y - \mathbf{Q}_y)}{\mathbf{J}_y - \mathbf{Q}_y}. \end{aligned} \tag{A.13}$$

With $\mathbf{N}_x$ determined, two more similar triangles calculate $\mathbf{N}_y$, one of which is the right angle triangle, $\triangle\mathbf{IDP}$, as shown in Figure A.6. Alternatively, $\mathbf{N}_y$ can be evaluated using the light ray, $\mathscr{L}_l$ (Equation A.4), since $\mathscr{L}_l \cdot \mathbf{N} = 0$. We get

$$\mathbf{N}_y(\mathbf{I}_x - \mathbf{P}_x) = \mathbf{N}_x(\mathbf{I}_y - \mathbf{P}_y) + \mathbf{I}_x\mathbf{P}_y - \mathbf{I}_y\mathbf{P}_x. \tag{A.14}$$

Substituting $\mathbf{N}_x$, from Equation A.13 into Equation A.14, reduces to

$$\mathbf{N}_y = \frac{\mathbf{P}_y(\mathbf{J}_y - \mathbf{M}_y) + \mathbf{I}_y(\mathbf{M}_y - \mathbf{Q}_y)}{\mathbf{J}_y - \mathbf{Q}_y}.$$

Therefore, the projection of a shadowing point $\mathbf{P}$ attached at $\mathbf{Q}$ to its shadowed point $\mathbf{N}$, on a parallel plane attached at $\mathbf{M}$ is

$$\begin{pmatrix} \mathbf{N}_x \\ \mathbf{N}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{J}_y - \mathbf{M}_y & 0 & \mathbf{I}_x(\mathbf{M}_y - \mathbf{Q}_y) \\ 0 & \mathbf{J}_y - \mathbf{M}_y & \mathbf{I}_y(\mathbf{M}_y - \mathbf{Q}_y) \\ 0 & 0 & \mathbf{J}_y - \mathbf{Q}_y \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix}.$$

As expected, the mapping is affine, the shadowed is a scaled version of the shadowing.

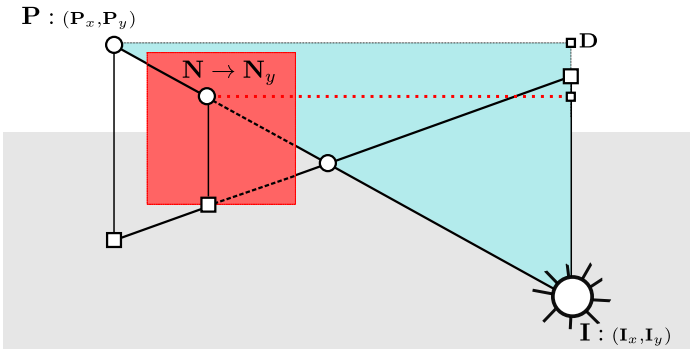


Figure A.6: Geometry for determining the height position, $\mathbf{N}_y$: the intersection of the light ray **PI** with the vertical axis **MN**.

This mapping produces the shadows in Figure 5.24. A shadow mask is scaled down to the proportions evaluated by the projection of the panel's boundary corners and then is placed in the appropriate corner in the perspective $\mathbb{RCS}$. In the rendering process this shadowed mask is painted using a clip to assure the projected shadow is always in contact with an opaque surface of the picture.

A.4.4 Comments

The shadow derivations demonstrate two points.

- There is room for generalization, which is done gradually as geometric patterns appear on the $\mathbb{RCS}$.
- The 2D geometry on the perspective $\mathbb{RCS}$ makes it possible to simulate the double projections that are frequent in nature. For example light projects illumination and shadows onto surfaces of 3D volumes that are then projected onto the image plane.

In this section on shadows, a simple projection onto the ground plane was generalized into shadows on the perspective $\mathbb{RCS}$ slices. Then, another type of shadow, one that is a scaled version of the shadowing object was analyzed. The mapping is simple but the context in which it is derived is interesting. A light behind the image plane and before the shadowing plane produces the same type of shadow. The derivation is similar, the matrix identical, because the geometry is equivalent. However, the similar triangles $\triangle$**PQS** and $\triangle$**JIS** are organized differently. The pseudo-illumination used for the derivation creates a geometric relationship between the 3D geometry and its 2D equivalent somewhat surprising because the geometry

is inverted. Yet it emerges from 2D geometry on the perspective $\mathbb{RCS}$, demonstrating the versatility of the representation.

Furthermore, the shadow mappings demonstrates that the physical $\mathbb{RCS}$ can vanish: the light vanishing point and attachment points within the perspective picture are sufficient to produce shadow effects. As artists who are comfortable to have the effect on the picture lead, my formalism computations become unconcerned of the physical implications: shadow effects are self-sufficiently generated within the perspective $\mathbb{RCS}$; not needing even the view parameters for the ground plane shadows.

A.5 Summary

When working directly on image plane, even more so in the perspective picture, with the formalism the perspective parameters are part of each computation, rather than being an afterthought. Continuous input from user is easily accepted and integrated in the 2D appearance: unlike a photograph, which projects to a 2D image only after objects have been in 3D and then cannot be modified in the picture appearance trivially.

Chapter 6 described the basic matrices used in developing an algebraic representation of the constructions described in Chapter 5. But it may give the impression of a user modelling in a physical 3D space then using the transformation to create the perspective picture. Nothing could be further from the truth; the manipulations computed in this appendix transform picture elements with the perspective $\mathbb{RCS}$, implementing the affordance available in my framework whose environment is 2D. They are the counterparts of the basic transformation—translating, rotating, scaling, texture mapping—of 3D modelling, which are compounded into the geometry that is rendered. Sets of affordances actually used will, of course, depend on the specific application.

The set of transformations developed in this appendix demonstrates the centrality of the ground plane and attachment points in making the 2D formalism work. The ground ray and the light ray intersect where the light ray encounters the opaque ground plane and can travel no farther. The preservation of cross ratios makes the constructions and derivations possible, regardless of the surprising layout of elements take after projection. While *lines map to lines*, there is an art in making them intersect to maintain invertible relations and stay perceptually effective. Artists have identified specific configurations and used them to good effect. It is important to bring them into computer environments, not hidden but to the surface, especially when making sense of complex geometry.

Appendix B

3D Mathematical Equivalence

Chapter 6 derives the 2D homogeneous matrices that reproduce the geometric constructions used by artists. This appendix calculates in 3D the image plane geometry of Renaissance perspective, proving that the artists' constructions are optically accurate.

B.1 Perspective

This section derives the model to image mappings, calculating in world coordinates, as seen in Figure B.1. The origin, **O**, lies at the centre of the canvas baseline; the x-axis is horizontal increasing to the right; the y-axis points up, and the z-axis points out of the canvas toward the viewer. The viewpoint, representing the viewer eye position is at $\mathbf{E} = (\mathbf{E}_x, \mathbf{E}_h, \mathbf{E}_d, 1)$ in the world frame, it is aligned to the world coordinate origin so that $\mathbf{E}_x = 0$.

B.1.1 Ground Plane

The goal is to derive the two-dimensional matrix that transforms a ground point, $\mathbf{q}'$ behind the canvas to its position **Q** on the perspective picture. First, consider a general point in the three-dimensional space behind the canvas, which is projected onto the image plane through the center of projection **E**: $\mathbf{r}'$ in Figure B.1 for example. The line through **E** and $\mathbf{r}'$ is $\mathbf{L}(t) = \mathbf{E} + t(\mathbf{r}' - \mathbf{E})$; solved for t when $\mathbf{L}(t) = \mathbf{R}$, gives $t = -\mathbf{E}_d/(\mathbf{r}'_z - \mathbf{E}_d)$ since $\mathbf{R}_z = 0$ by definition. Thus, we get

$$\mathbf{R}_x = \frac{\mathbf{E}_x\mathbf{r}'_z - \mathbf{E}_d\mathbf{r}'_x}{\mathbf{r}'_z - \mathbf{E}_d} \qquad \text{and} \qquad \mathbf{R}_y = \frac{\mathbf{E}_h\mathbf{r}'_z - \mathbf{E}_d\mathbf{r}'_y}{\mathbf{r}'_z - \mathbf{E}_d}.$$

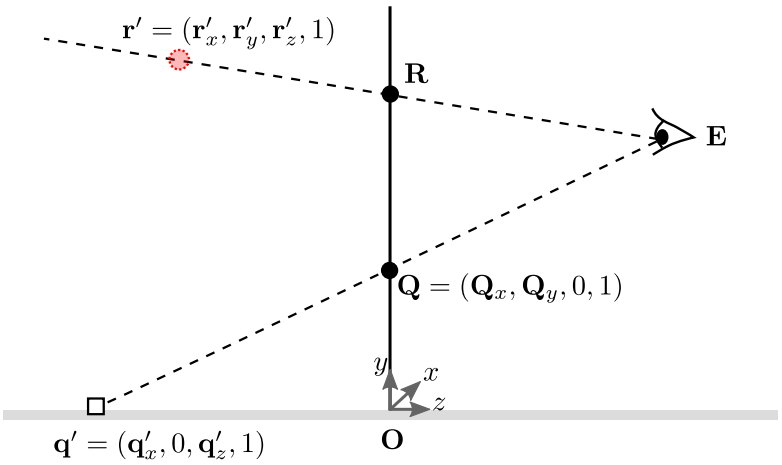


Figure B.1: The projective geometry of points in 3D, mapped to their images on the canvas is shown in an orthogonally-projected side view. **E** is the viewpoint; $\mathbf{r}'$ is a general point in 3D. $\mathbf{q}'$ is a point on the ground plane. **R** and **Q** are the projections onto the canvas of $\mathbf{r}'$ and $\mathbf{q}'$ respectively.

The projection of a general point $\mathbf{r}'$ on the line of sight to its corresponding point on the canvas **R** is written in terms of a homogeneous matrix as follows,

$$\begin{pmatrix} \mathbf{R}_x \\ \mathbf{R}_y \\ \mathbf{R}_z \\ 1 \end{pmatrix} \sim \vartheta' \begin{pmatrix} \mathbf{r}'_x \\ \mathbf{r}'_y \\ \mathbf{r}'_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & \mathbf{E}_x & 0 \\ 0 & -\mathbf{E}_d & \mathbf{E}_h & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{r}'_x \\ \mathbf{r}'_y \\ \mathbf{r}'_z \\ 1 \end{pmatrix}. \tag{B.1}$$

The matrix ϑ' is not invertible as it reduces the dimensions of any point in 3D to its point on the canvas: all the points on the ray $\mathbf{E}\mathbf{r}'$ maps to **R**. In contrast, when we consider a 3D point on the ground plane behind the canvas its height coordinate is fixed, that is, $\mathbf{q}'$ in Figure B.1 has $\mathbf{q}'_y = 0$. Then, a row and column in ϑ' are unnecessary and can be removed. Thus, since by definition $\mathbf{E}_x = 0$, we get the following two-dimensional mapping transforming the depth coordinate of a point $\mathbf{q}'$ on the ground plane, to the height coordinate of **Q** on the canvas.

$$\begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h & 0 \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{q}'_x \\ \mathbf{q}'_z \\ 1 \end{pmatrix}. \tag{B.2}$$

This matrix is invertible and corresponds to ϖ (Equation 6.1), the 2D homogeneous matrix that reproduces the canvas geometric constructions used by artists.

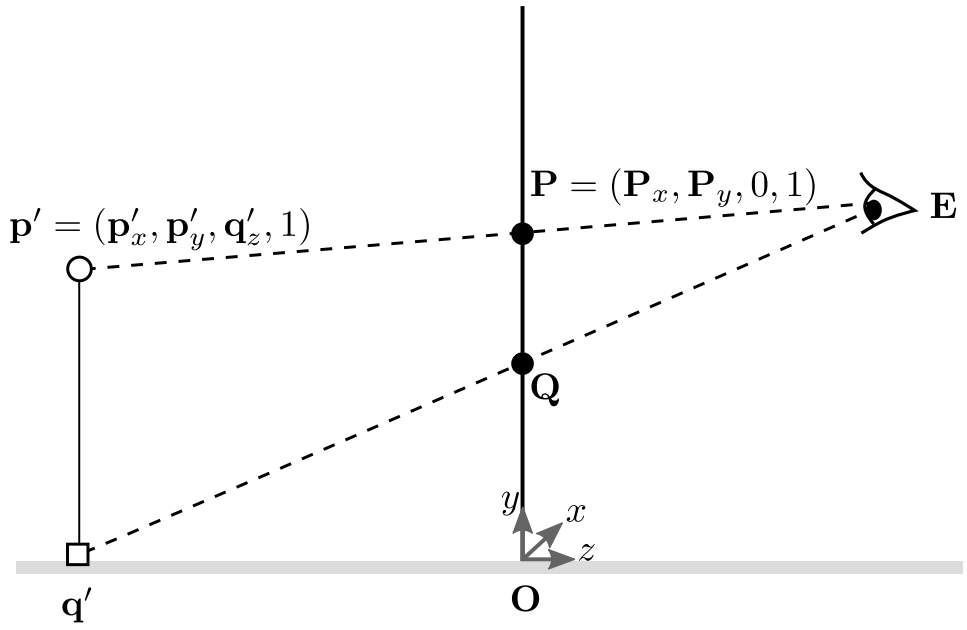


Figure B.2: The projection geometry for the points in depth and their images on the canvas is shown in side view. $\mathbf{E}$ is the viewpoint; $\mathbf{q}'$ is a point on the ground plane; $\mathbf{p}'$ is a general point attached at $\mathbf{q}'$. $\mathbf{Q}$ and $\mathbf{P}$ are the projections on the canvas of $\mathbf{q}'$ and $\mathbf{p}'$ respectively, through the center of projection $\mathbf{E}$.

B.1.2 Above the Ground Plane

To transform a general 3D point to the canvas, consider the general point, $\mathbf{p}'$, with its corresponding attachment point, $\mathbf{q}'$, and their projections on the canvas, $\mathbf{P}$ and $\mathbf{Q}$ respectively, as shown in Figure B.2. Solving the lines $\mathbf{L}_{\mathbf{p}'}(s) = \mathbf{E} + s(\mathbf{p}' - \mathbf{E})$ and $\mathbf{L}_{\mathbf{q}'}(t) = \mathbf{E} + t(\mathbf{q}' - \mathbf{E})$ for $\mathbf{L}_{\mathbf{p}'}(s) = \mathbf{P}$ and $\mathbf{L}_{\mathbf{q}'}(t) = \mathbf{Q}$ we get $s = t$ as both $\mathbf{P}_z = \mathbf{Q}_z = 0$, and $\mathbf{p}'_z = \mathbf{q}'_z$. Thus, it can be solved as in the previous section,

$$\mathbf{P}_x = \mathbf{Q}_x = \frac{\mathbf{E}_x\mathbf{q}'_z - \mathbf{E}_d\mathbf{p}'_x}{\mathbf{q}'_z - \mathbf{E}_d}, \quad \mathbf{P}_y = \frac{\mathbf{E}_h\mathbf{q}'_z - \mathbf{E}_d\mathbf{p}'_y}{\mathbf{q}'_z - \mathbf{E}_d} \quad \text{and } \mathbf{Q}_y = \frac{\mathbf{E}_h\mathbf{q}'_z}{\mathbf{q}'_z - \mathbf{E}_d},$$

which, inverting the mappings, can be written as

$$\mathbf{p}'_x = \mathbf{q}'_x = \frac{\mathbf{E}_x\mathbf{Q}_y - \mathbf{E}_h\mathbf{P}_x}{\mathbf{Q}_y - \mathbf{E}_h}, \quad \mathbf{p}'_y = \frac{\mathbf{E}_h\mathbf{Q}_y - \mathbf{E}_h\mathbf{P}_y}{\mathbf{Q}_y - \mathbf{E}_h} \quad \text{and } \mathbf{p}'_z = \frac{\mathbf{E}_d\mathbf{Q}_y}{\mathbf{Q}_y - \mathbf{E}_h}.$$

This is a projective transformation and is written with the homogeneous matrix, $(\pi')^{-1}$ as

$$\begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{p}'_z \\ 1 \end{pmatrix} \sim (\pi')^{-1} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_h & 0 & \mathbf{E}_x & 0 \\ 0 & -\mathbf{E}_h & \mathbf{E}_h & 0 \\ 0 & 0 & \mathbf{E}_d & 0 \\ 0 & 0 & 1 & -\mathbf{E}_h \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ \mathbf{Q}_y \\ 1 \end{pmatrix}, \tag{B.3}$$

which can be inverted and simplified as

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ \mathbf{Q}_y \\ 1 \end{pmatrix} \sim \pi' \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{q}'_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & \mathbf{E}_x & 0 \\ 0 & -\mathbf{E}_d & \mathbf{E}_h & 0 \\ 0 & 0 & \mathbf{E}_h & 0 \\ 0 & 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{q}'_z \\ 1 \end{pmatrix}$$
$$\sim \begin{pmatrix} -\mathbf{E}_d & 0 & \mathbf{E}_x & 0 \\ 0 & 0 & \mathbf{E}_h & -\mathbf{p}'_y \mathbf{E}_d \\ 0 & 0 & \mathbf{E}_h & 0 \\ 0 & 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{p}'_x \\ 0 \\ \mathbf{q}'_z \\ 1 \end{pmatrix}.$$

Since $\mathbf{E}_x = 0$ the matrix can be summarized by a two-dimensional mapping, solving the location on the canvas given the point placement in depth on the ground plane and its height, $\mathbf{p}'_y$, behind the canvas. Indeed, the following 2D matrix moves a 3D general point, $\mathbf{p}'$, attached on the ground plane at $\mathbf{q}'$ to its location $\mathbf{P}$ on the canvas:

$$\begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & 0 \\ 0 & \mathbf{E}_h & -\mathbf{p}'_y \mathbf{E}_d \\ 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{q}'_z \\ 1 \end{pmatrix}. \tag{B.4}$$

This 2D matrix transforms a point in depth on the physical ground plane at height $\mathbf{p}'_y$ to a 2D point, $\mathbf{P}$ on the canvas: it is the 3D equivalent to the matrix $\pi_{\mathbf{c}'}$ (Equation 6.3), where $\mathbf{c}' = \mathbf{p}'_y$, and is invertible.

B.2 Shadows

This section demonstrates the optical correctness of the matrix Ψ (Equation A.5), which evaluates on the canvas the projection onto the ground plane of a general shadowing point with respect to the illumination position. Figure B.3 shows the 3D geometry in a side view, including the illumination and shadow point positions in 3D and their projections onto the canvas.

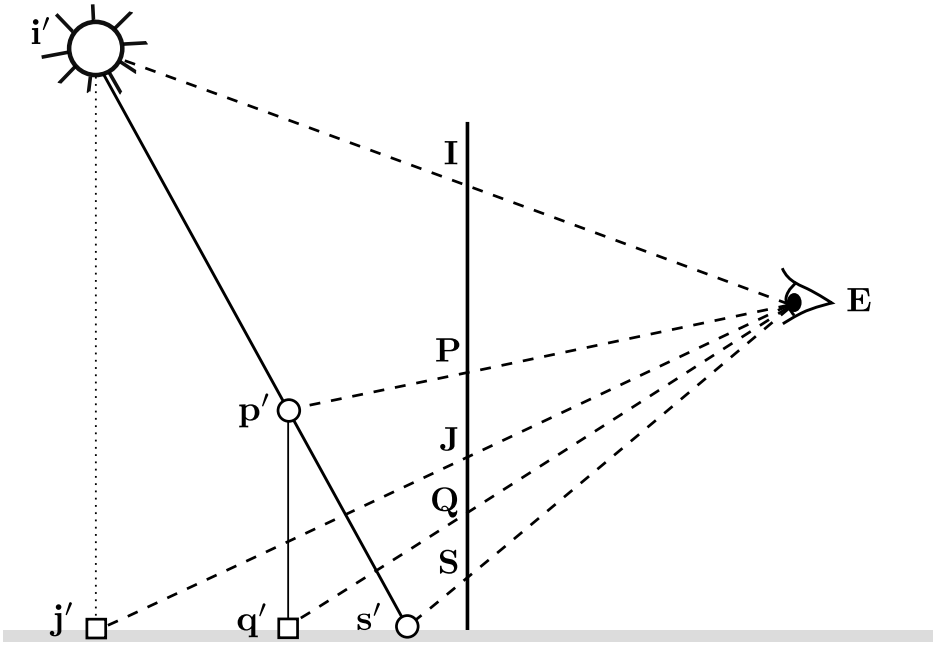


Figure B.3: The shadow casting geometry is shown in side view. $\mathbf{i}'$ is the light source; $\mathbf{j}'$ is its attachment point. $\mathbf{p}'$ is the shadowing point, $\mathbf{q}'$ its attachment point. $\mathbf{s}'$ is the shadowed location on the ground plane. These points are projected to the image plane with the eye $\mathbf{E}$ as the centre of projection. Their images have the same unprimed label in upper-case.

The source of illumination in 3D is at $\mathbf{i}' = (\mathbf{i}'_x, \mathbf{i}'_y, \mathbf{i}'_z, 1)$, which may be before the image plane ($\mathbf{i}'_z > 0$) or behind it ($\mathbf{i}'_z < 0$). The shadowing point $\mathbf{p}' = (\mathbf{p}'_x, \mathbf{p}'_y, \mathbf{p}'_z, 1)$ that casts the shadow is behind the image plane. Both $\mathbf{i}'$ and $\mathbf{p}'$ have accompanying attachment points: the attachment point of the source of illumination is the light vanishing point, $\mathbf{j}' = (\mathbf{i}'_x, 0, \mathbf{i}'_z, 1)$; the attachment point of the shadowing point is $\mathbf{q}' = (\mathbf{p}'_x, 0, \mathbf{p}'_z, 1)$. The shadow is cast on the 3D ground plane, the xz-plane, at the shadowed point $\mathbf{s}'$. Each of these five points has a projection on the canvas.

When placing the shadow the artist already knows the positions on the canvas of the source of illumination **I**, of the shadowing point, **P** and of their two attachment points, **J** and **Q**. Based on the positions of these points, the artist places on the canvas construction lines to determine the shadowed point **S** (§ 5.5.3). Calculating explicitly the position **S** in terms of the other four canvas points shows how to create the shadow-casting geometry on the image plane: the mapping for **S** using the 3D points $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{p}'$ and $\mathbf{q}'$, demonstrates that the 2D matrix derived for the formalism of Section A.4.1 is optically correct.

B.2.1 On the 3D Ground Plane

Figure B.3 shows the 3D shadowing geometry for the simplest shadowed surface, the ground plane: the light source, $\mathbf{i}'$, throws a ray through a shadowing point, $\mathbf{p}'$, that casts the ground shadowed point $\mathbf{s}'$. The light ray $\mathbf{L}(t) = \mathbf{p}' + t(\mathbf{i}' - \mathbf{p}')$ intersects the shadowed plane at $y = 0$. Solving for $\mathbf{L}(t) = \mathbf{s}'$ we get $t = \mathbf{p}'_y / (\mathbf{p}'_y - \mathbf{i}'_y)$ and thus

$$\mathbf{s}'_x = \frac{\mathbf{i}'_x \mathbf{p}'_y - \mathbf{i}'_y \mathbf{p}'_x}{\mathbf{p}'_y - \mathbf{i}'_y} \qquad \text{and} \qquad \mathbf{s}'_z = \frac{\mathbf{i}'_z \mathbf{p}'_y - \mathbf{i}'_y \mathbf{p}'_z}{\mathbf{p}'_y - \mathbf{i}'_y}.$$

Where the shadow falls on the 3D ground plane can be expressed in terms of the 3D homogeneous matrix, Σ, as

$$\begin{pmatrix} \mathbf{s}'_x \\ \mathbf{s}'_y \\ \mathbf{s}'_z \\ 1 \end{pmatrix} \sim \Sigma \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{p}'_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{i}'_y & \mathbf{i}'_x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{i}'_z & -\mathbf{i}'_y & 0 \\ 0 & 1 & 0 & -\mathbf{i}'_y \end{pmatrix} \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{p}'_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{p}'_y & -\mathbf{p}'_x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\mathbf{p}'_z & \mathbf{p}'_y & 0 \\ 0 & -1 & 0 & -\mathbf{p}'_y \end{pmatrix} \begin{pmatrix} \mathbf{i}'_x \\ \mathbf{i}'_y \\ \mathbf{i}'_z \\ 1 \end{pmatrix},$$

where the row of zeros occurs because the shadow is cast on the ground plane, i.e., $\mathbf{s}'_y = 0$. The two separate expressions indicate that it is possible to regard the creation of the shadow as the illumination projecting the shadowing point, or as the shadowing point projecting the source of illumination.

B.2.2 On the Canvas Ground Plane

The shadowed point $\mathbf{s}'$ is projected to the image plane using the matrix ϑ' (Equation B.1), specifying the canvas location of the shadowed point $\mathbf{S}$. Thus, for a shadowing point, $\mathbf{p}'$, behind the image plane, $\mathbf{S}$ is calculated as $\mathbf{S} \sim \vartheta' \Sigma \mathbf{p}'$:

$$\begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \\ \mathbf{S}_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\mathbf{E}_d & 0 & \mathbf{E}_x & 0 \\ 0 & -\mathbf{E}_d & \mathbf{E}_h & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\mathbf{E}_d \end{pmatrix} \begin{pmatrix} -\mathbf{i}'_y & \mathbf{i}'_x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{i}'_z & -\mathbf{i}'_y & 0 \\ 0 & 1 & 0 & -\mathbf{i}'_y \end{pmatrix} \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{p}'_z \\ 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} \mathbf{E}_d \mathbf{i}'_y & \mathbf{E}_x \mathbf{i}'_z - \mathbf{E}_d \mathbf{i}'_x & -\mathbf{E}_x \mathbf{i}'_y & 0 \\ 0 & \mathbf{E}_h \mathbf{i}'_z & -\mathbf{E}_h \mathbf{i}'_y & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{i}'_z - \mathbf{E}_d & -\mathbf{i}'_y & \mathbf{E}_d \mathbf{i}'_y \end{pmatrix} \begin{pmatrix} \mathbf{p}'_x \\ \mathbf{p}'_y \\ \mathbf{p}'_z \\ 1 \end{pmatrix},$$

which extends to the following relation for a canvas point $\mathbf{P}$ using $(\pi')^{-1}$ (Equation B.3)

$$\begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \\ \mathbf{S}_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{E}_d\mathbf{i}'_y & \mathbf{E}_x\mathbf{i}'_z - \mathbf{E}_d\mathbf{i}'_x & -\mathbf{E}_x\mathbf{i}'_y & 0 \\ 0 & \mathbf{E}_h\mathbf{i}'_z & -\mathbf{E}_h\mathbf{i}'_y & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \mathbf{i}'_z - \mathbf{E}_d & -\mathbf{i}'_y & \mathbf{E}_d\mathbf{i}'_y \end{pmatrix} \begin{pmatrix} -\mathbf{E}_h & 0 & \mathbf{E}_x & 0 \\ 0 & -\mathbf{E}_h & \mathbf{E}_h & 0 \\ 0 & 0 & \mathbf{E}_d & 0 \\ 0 & 0 & 1 & -\mathbf{E}_h \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ \mathbf{Q}_y \\ 1 \end{pmatrix}$$

$$\sim \mathbf{E}_h \begin{pmatrix} -\mathbf{E}_d\mathbf{i}'_y & \mathbf{E}_d\mathbf{i}'_x - \mathbf{E}_x\mathbf{i}'_z & -(\mathbf{E}_d\mathbf{i}'_x - \mathbf{E}_x\mathbf{i}'_z) & 0 \\ 0 & -\mathbf{E}_h\mathbf{i}'_z & \mathbf{E}_h\mathbf{i}'_z - \mathbf{E}_d\mathbf{i}'_y & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -(\mathbf{i}'_z - \mathbf{E}_d) & \mathbf{i}'_z - \mathbf{E}_d & -\mathbf{E}_d\mathbf{i}'_y \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ \mathbf{Q}_y \\ 1 \end{pmatrix}.$$

Using $(\pi')^{-1}$ we also have the following relations for $\mathbf{i}'$ and $\mathbf{j}'$

$$\mathbf{i}'_x = \frac{\mathbf{E}_x\mathbf{J}_y - \mathbf{E}_h\mathbf{I}_x}{\mathbf{J}_y - \mathbf{E}_h}, \quad \mathbf{i}'_y = \frac{\mathbf{E}_h\mathbf{J}_y - \mathbf{E}_h\mathbf{I}_y}{\mathbf{J}_y - \mathbf{E}_h} \quad \text{and} \quad \mathbf{i}'_z = \frac{\mathbf{E}_d\mathbf{J}_y}{\mathbf{J}_y - \mathbf{E}_h}.$$

Thus

$$\begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \\ \mathbf{S}_z \\ 1 \end{pmatrix} \sim \frac{\mathbf{J}_y - \mathbf{E}_h}{\mathbf{E}_d} \begin{pmatrix} \mathbf{I}_y - \mathbf{J}_y & -\mathbf{I}_x & \mathbf{I}_x & 0 \\ 0 & -\mathbf{J}_y & \mathbf{I}_y & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & \mathbf{I}_y - \mathbf{J}_y \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ \mathbf{Q}_y \\ 1 \end{pmatrix},$$

which is equivalent to

$$\begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \\ \mathbf{S}_z \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{I}_y - \mathbf{J}_y & -\mathbf{I}_x & 0 & \mathbf{I}_x\mathbf{Q}_y \\ 0 & -\mathbf{J}_y & 0 & \mathbf{I}_y\mathbf{Q}_y \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & \mathbf{I}_y - \mathbf{J}_y + \mathbf{Q}_y \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 0 \\ 1 \end{pmatrix}.$$

This matrix shows that ground plane shadows can be calculated using only canvas points. Furthermore, since $\mathbf{S}_z = 0$ on the image plane we obtain the following 2D matrix

$$\begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{I}_y - \mathbf{J}_y & -\mathbf{I}_x & \mathbf{I}_x\mathbf{Q}_y \\ 0 & -\mathbf{J}_y & \mathbf{I}_y\mathbf{Q}_y \\ 0 & -1 & \mathbf{I}_y - \mathbf{J}_y + \mathbf{Q}_y \end{pmatrix} \begin{pmatrix} \mathbf{P}_x \\ \mathbf{P}_y \\ 1 \end{pmatrix},$$

which is the same as the matrix Ψ (Equation A.5). This section showed that the artists' construction for ground plane shadows on the canvas is optically correct in terms of 3D geometry.

Appendix C

Projective Geometry

Based on the tiled floor distance point construction of the Renaissance (Chapter 4), we have seen that the combination of the pvp and the dp serves to represent the 3D eye on the canvas. The $\mathbb{RCS}$ mappings in Chapter 6 that derive the perspective picture are a function of $\mathbf{E}_h$ and $\mathbf{E}_d$, the dp coordinates on the canvas, which express the 3D observation point out of the picture. The question that follows is: is it the only 3D eye representation on the canvas? Or is it just one useful representation? In Piero's diagonal construction, there is a fold of the physical floor on the image plane, the dp is not where the 3D eye point would move to if the full ground was rotated rather than the half-plane: the dp is only really a convergence point on the horizon of the floor diagonals. This appendix discusses another eye's that exists in the $\mathbb{RCS}$.

Renaissance artists practised perspective. I argue in this thesis that the concordance between the tiled floor representations, the physical one and the perspective one, was fundamental to understand and develop the simulation of 3D directly onto the canvas. Appendix B shows that their representations are optically correct in terms of the 3D geometry simulated. Renaissance artists created 3D projections of the world onto the canvas, without having formalized the general geometry: projective geometry. Projective geometry maps shapes between spaces. This includes for example the mapping of a volume or a flat shape onto its image on a given plane, as well as the mapping between two different images of a same shape.

Indeed, Renaissance practices, without knowing it, are conceptual applications of projective geometry, the theory of which was only described later: they used properties of the projective space (the 3D one and 2D one) while they were not yet formally established. Bridging between art and science, mathematicians since the Renaissance worked on the formalization of the geometry involved in the application of perspective. In 1822, Poncelet published the foundational

treatise on projective geometry. But long before then significant results were achieved, as this appendix describes. In particular Gérard Desargues, (1591–1661), who lived two centuries before Poncelet, is considered one of the founder of projective geometry.

Desargues theorem relates images of two corresponding shapes and the invariances that follow by the principle of projection: there exist an axis of projection and a center. While the Renaissance practice of juxtaposing two images of the floor representation makes the invariant axis clear, the baseline, the invariant center on the canvas plane was neither highlighted, nor used. The goal of this appendix is to discuss this important point and its link to the distance point. In effect, the dp is a convenient and ingenious substitute for the projection center, which is sometimes called *the auxiliary eye point* or *the turned-in eye point* [4, p. 343]. So much so that while interesting for the theorization of the geometry occurring, it was unnecessary for its practice as the dp fulfills the same role: stipulating the distance of the viewer from the canvas.

For completeness, this appendix gives a taste of the connection between these two fundamental points on the canvas plane: the projective geometry auxiliary eye point, labelled V in this appendix; and the distance point, dp, used in Renaissance construction methods. The constructive geometry presented here was inspired by Coxeter [21]. For further information about the turned-in eye point, Andersen's book should be consulted [4]. This appendix does not present formal proofs, but rather demonstrations by examples, aiming to satisfy intuition but falling short of convincing mathematical minds.

C.1 Connection with Desargues Theorem

I first present Desargues theorem and examples of its appearance in the 2D and 3D projective space. Then I demonstrate that Desargues theorem appears in the constructive space of my canvas where the physical representation of the tiled floor and the perspective one are juxtaposed. Doing so reveals a special invariant point, which is both related to the 3D eye point and to the distant point.

Desargues' theorem is as follows:

> *In a projective space, two triangles are in perspective from a point if and only if they are in perspective from a line.*

That is, two triangles, $\triangle ABC$, $\triangle A'B'C'$, not necessarily in the same plane, are in perspective according to a *the center of perspectivity*, V, if and only if the intersections of the three pairs of side' AB, $A'B'$; BC, $B'C'$; and AC, $A'C'$ are all points on a line, which is called *the axis of perspectivity*.

C.1.1 Desargues in 2D

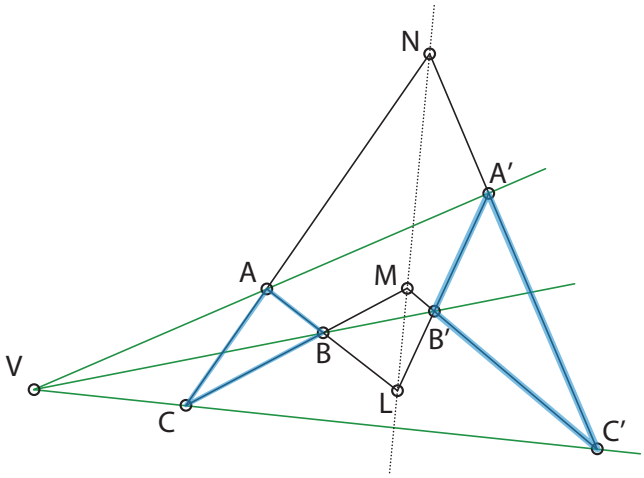


Figure C.1: Illustration of a Desargues theorem triangles configuration in the plane. Two triangles, $\triangle ABC$, $\triangle A'B'C'$, in perspective according to the center of perspectivity, V, which is where AA', BB' and CC' concurrent, and where the intersections L, M and N, of AB, $A'B'$; BC, $B'C'$; and AC, $A'C'$ respectively, are collinear.

A plane is a 2D projective space. Figure C.1 shows an instance of Desargues theorem in the 2D projective space: two triangles on the same plane where both a center of perspectivity, V, and a perspectivity axis exist. (While it should be simple, Desargues theorem does not work in general for 2D projective space—the center does not guarantee the axis, additional constraints are needed— but it is good illustration to understand the 3-space version.) How to obtain the figures of two triangles showing Desargues' theorem? Consider any triangle $A'B'C'$ and any line, not parallel to any of the triangle edges, to be the perspectivity axis. The triangle edges $A'B'$, $B'C'$ and $A'C'$ intersect the perspectivity axis at L, M and N. Now, select a perspectivity center, V, and draw the rays, VA', VB' and VC'. Construct one of the rays corresponding to a vertex of the triangle $\triangle ABC$, by placing either A on VA' or B on VB' or C on VC' along it. Then place the second vertex of $\triangle ABC$ on its ray respecting the intersection with the axis. Do the same for the last vertex. The edge between the first placed vertex and the last one, intersects at the axis unused point.

The 2D space illustrates the concepts of the theorem, but it is hard to relate it to the real world phenomena of image creation.

C.1.2 Desargues in 3D

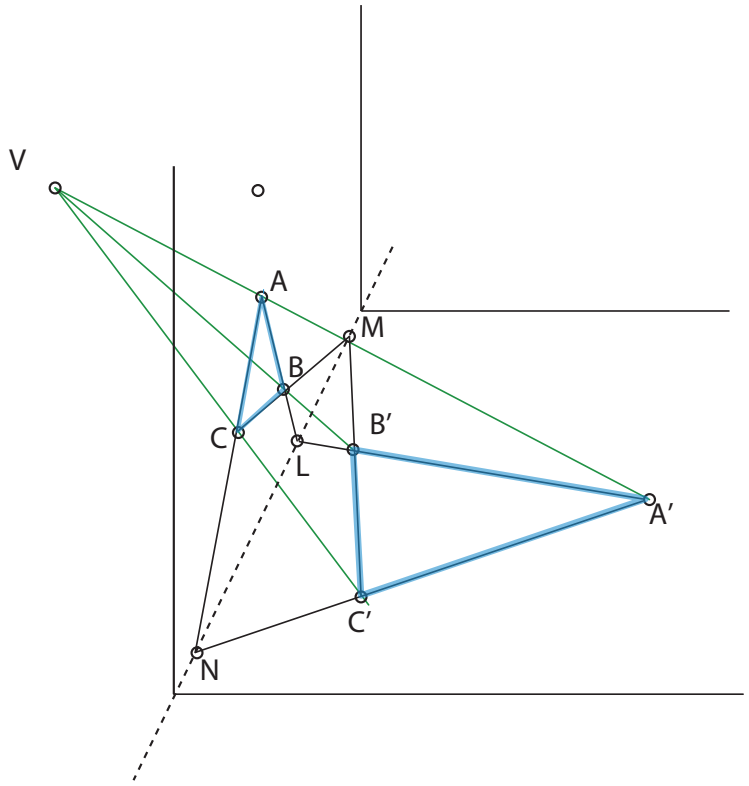


Figure C.2: Desargues theorem in the physical space: the two triangles are onto distinct planes. One is on the physical ground, while the other is on the image plane according to the center of perspectivity, V, corresponding to the eye. The lines of the triangle edges meet at the baseline.

Therefore, consider another example of Desargues theorem, illustrated in Figure C.2, that holds in the 3D projective space: two triangles with Desargues relations that are not on the same plane. In fact, the two planes used are positioned in the most common configuration for creating an image, where a shape on the physical plane is to be represented on the image plane. Given the location of the 3D eye, V, and the triangle on the physical plane, $\triangle A'B'C'$, it is possible to derive its image on the image plane. The lines that are the rays VA', VB' and VC' intersect the image plane at A, B and C respectively. Doing so creates the triangle on the image plane ABC which is in Desargues configuration with $\triangle A'B'C'$, where V is the center of perspectivity. Therefore, according to Desargues theorem there also exists an axis of perspectivity. Indeed, the triangle edge pairs in both planes, which are AB, $A'B'$; BC, $B'C'$;

and AC, $A'C'$ produce common points L, M and N that are collinear and lie on the line where the two planes involved intersect: the baseline, the line that we have seen is fundamental to transfer measures from the physical space to the perspective picture.

In effect, the geometric constructions of this thesis rely on this special property of the baseline. Given a line on the physical plane, we know that its perspective representation joins at the baseline. While there are many representations possible if the viewer distance to the image plane is not known, one aspect is always determined: the baseline connection. For example, take any floor orthogonal on the physical plane, we know that it meets at the baseline with its representation on the perspective space, even if we know nothing else. This invariance applies for any line orientations, with the exception of lines that are parallel to the baseline as they do not intersect the baseline and where appearance demands the foreshortened relation to be specified.

If V is moved in Figure C.2, the image ABC of $A'B'C'$ changes on the image plane, but the baseline intersections L, M and N stay the same. If V is on the image plane then the triangle $\triangle ABC$ collapses to V.

C.1.3 Desargues in the $\mathbb{RCS}$

In this thesis, inspired by Piero's diagram, a special canvas space, the $\mathbb{RCS}$, was used to structure the construction of shapes in the perspective picture using direct construction lines on the image plane surface. In effect, it combines on the 2D projective space of the image plane, the perspective representation of a tiled floor with the representation of the physical floor of the 3D projective space, rotated so as to lie on the plane. Indeed, the physical ground plane of the 3D projective space was folded—called *a rabatment* by Andersen [4, p. xxxii]—to be on the 2D projective space of the canvas. Doing so brings to the image plane the physical floor, producing a 2D projective plane where the two half planes each contains a different image space.

If we consider the example above, it means that the triangle $\triangle A'B'C'$ of Figure C.2 is brought onto the image plane of $\triangle ABC$. Are then the two triangles in a Desargues relation as in Figure C.1? If yes, where is the perspectivity center? How does it relate to the eye, the 3D projective space V of Figure C.2?

Consider a triangle $\triangle A'B'C'$ on the rotated physical floor and its corresponding shape $\triangle ABC$ on the perspective floor: Figure C.3 shows the result of the construction of the perspective image of the triangle (it can either be done using the corners of tile correspondence or if the tiled floor was not apparent then by applying the vertex-by-vertex the algorithm described in § 5.3).

The baseline is the line where the two floor spaces coincide. For the two representations of the triangle, each pair of lines passing through their vertices indeed meets at a common point on the baseline, that is L, M and N, respectively. (This property of the baseline intersection space for the two representations of the same line has been used in this thesis; see Figures5.15 (b) or F.2.) This correspondence put the two triangles in the Desargues' theorem relation: in the common plane they are in perspective from a line, the baseline (as was the case in Figure C.2). Indeed, the property within the 3D projective space is unchanged by rotating the plane through that axis to collapse the 3D space to a planar one. From Desargues' theorem we know that the two triangles are in perspective from a point in the plane. But, where is this point? And to what does it correspond?

Could this point be the pvp? No, the pvp is not invariant between the physical $\mathbb{RCS}$ and perspective one. From the formalism presented in Chapter 6, we know that the pvp on the horizon on the perspective $\mathbb{RCS}$ maps to a point at infinity on the physical $\mathbb{RCS}$ floor of the image plane. Could it be dp? No, similarly the dp maps to infinity below the baseline in the physical $\mathbb{RCS}$.

Using Figure C.3, we show by example that the lines AA', BB', and CC' meet at a point above the horizon. We observe that this convergent point is vertically aligned with the pvp and that its distance above it is the same as the distance between pvp and dp. Projective geometry mathematicians, in particular 'sGravesande (1688–1742) and Ozanam (1640–c. 1717), have shown and used the location of this particular projection center on the canvas [4, p. 343].

Examining Figure C.2, we know that the plane containing the triangle $\triangle A'B'C'$ underwent a rotation of 90° around the baseline to be on the image plane in Figure C.3. Similarly, V in Figure C.2 should have undergone the same transformation, while $\triangle ABC$ stays fixed to obtain Figure C.3. The convergent center on the image plane in Figure C.3 is in effect an auxiliary eye point: the V in Figure C.2 rotated on the canvas. Concretely, the plane parallel to the ground containing V—which contains the line of sight and the horizon—is rotated 90° around the horizon so as to be on the image plane. Undergoing a simultaneous rotation, the rays from the rotated eye passing through A, B and C stay respectively aligned with A', B' and C' which are on the physical ground plane that also rotates. (This is related to the invariance of a perspective image under simultaneous rotations, discovered by Stevin (1548–1620) [4, p. 271].)

Doing so happens to precisely locate V, the auxiliary eye, on the canvas:

1. it is vertically above the principal vanishing point, and
2. at a vertical distance above it equal to the distance separating the pvp from the dp.

Thus, V on Figure C.3, is at $(0, \mathbf{E}_h + \mathbf{E}_d)$.

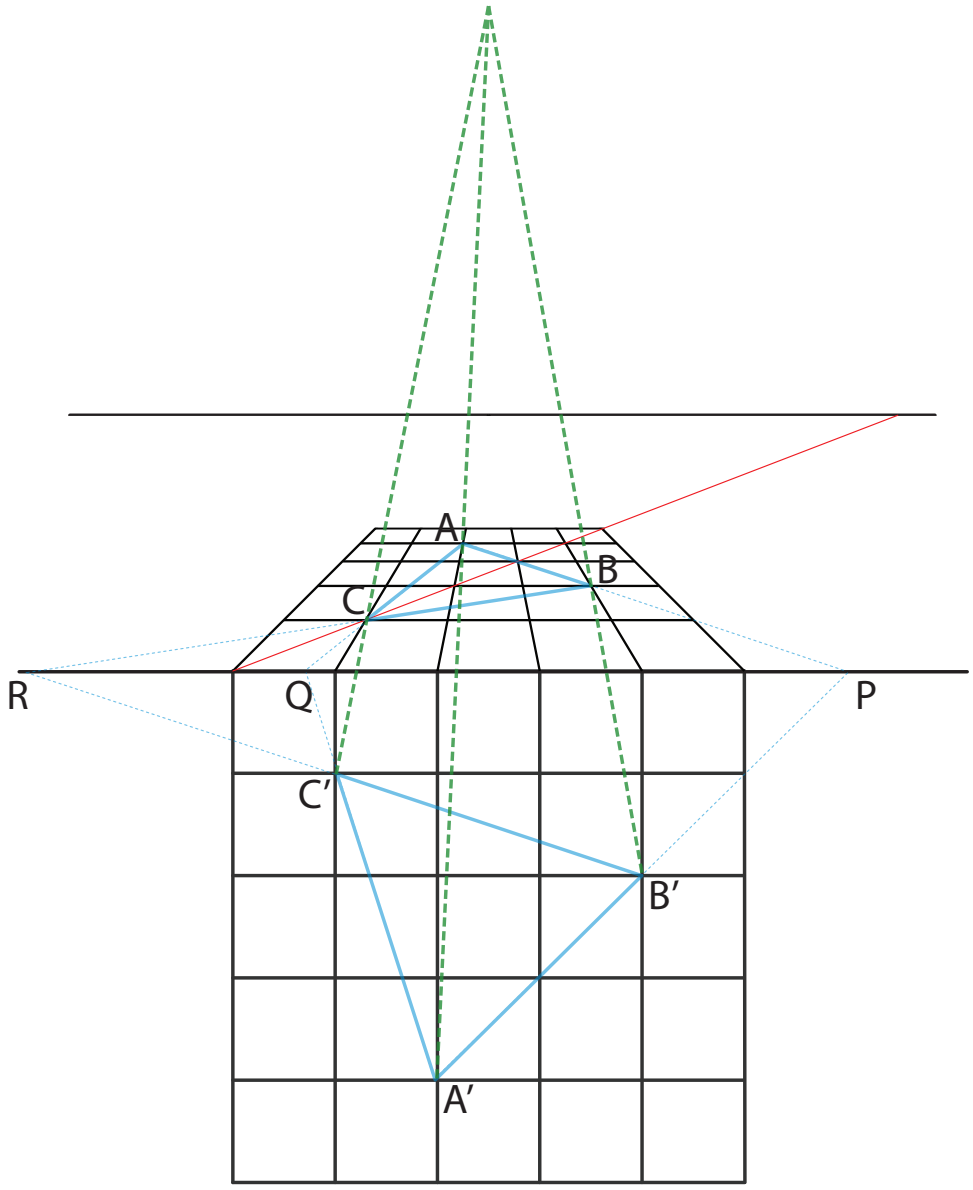


Figure C.3: Desargues theorem relating two corresponding triangles on each floor of the ℝℂ𝕊. $\triangle A'B'C'$ on the physical ℝℂ𝕊 floor corresponds to $\triangle ABC$ on the perspective one. Each can be constructed geometrically using Piero's diagonal. Rays from the auxiliary eye link the respective vertex of the triangles (the green dashed lines).

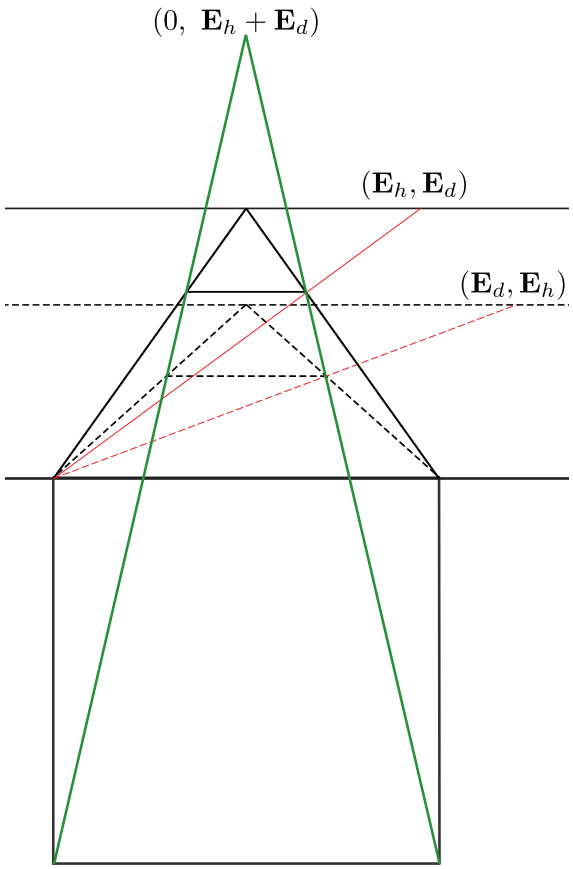


Figure C.4: A usual tiled floor in perspective constructed using a dp at $(\mathbf{E}_d, \mathbf{E}_h)$ (dashed lines) above the physical corresponding floor. A second tiled floor in perspective constructed using a dp with swapped coordinates, at $(\mathbf{E}_h, \mathbf{E}_d)$ (solid lines). The two tiled floors in perspective relate by rays from the auxiliary eye point at $(0, \mathbf{E}_h + \mathbf{E}_d)$, Desargues theorem center of Figure C.3: the two green lines pass through each of their extreme horizontal boundary in depth, as well as the corresponding vertex of the physical floor edge.

Thus, Figure C.4 is now not too surprising. For a 3D viewpoint that is instead at a height $\mathbf{E}_d$ and a distance $\mathbf{E}_h$ away from the image plane, the perspective tiled floor is constructed (solid lines): the same auxiliary point, $(0, \mathbf{E}_h + \mathbf{E}_d)$ as for the perspective tiled floor using the usual dp at $(\mathbf{E}_d, \mathbf{E}_h)$, holds. The auxiliary point is the result of combining two folds: one around the baseline, the other around the horizon. The figure illustrates the relation between the mapping ϖ (from the physical $\mathbb{R}$CS ground plane and the perspective one) and its inverse mapping, where the $\mathbf{E}_h$ and $\mathbf{E}_d$ coordinates are swapped (§ 6.2).

To understand deeper the connection, consider the following simple scenario on the ℝC𝕊, illustrated in Figure C.5 (a). We have a square physical floor, a baseline, and a principal vanishing point on the horizon: these elements determine the relative position of the viewer from the canvas and the scene, without specifying its distance, only where the viewer is facing. The appearance of the perspective floor changes according to the distance between the viewer and the image plane: the pvp does not encode this parameter, but both the distance point and the auxiliary eye point do. Figure C.5 (b) shows different perspective floor appearances for various dp positions, using the Piero's diagonal construction (§ 4.2.3). Similarly, Figure C.5 (c) shows the same different perspective floor appearances constructed using lines through the respective auxiliary eye points: the lines joining them to the physical floor extreme corner intersect on the same side the main floor orthogonal on the perspective space at the location where the floor respective areas ought to terminate. While the position of the auxiliary point has not been proven geometrically, it is clear that its position has a correspondence with the distance point: on the canvas it also encodes how far the viewer is from the scene but according to its vertical distance above pvp, rather than its horizontal distance as the dp does.

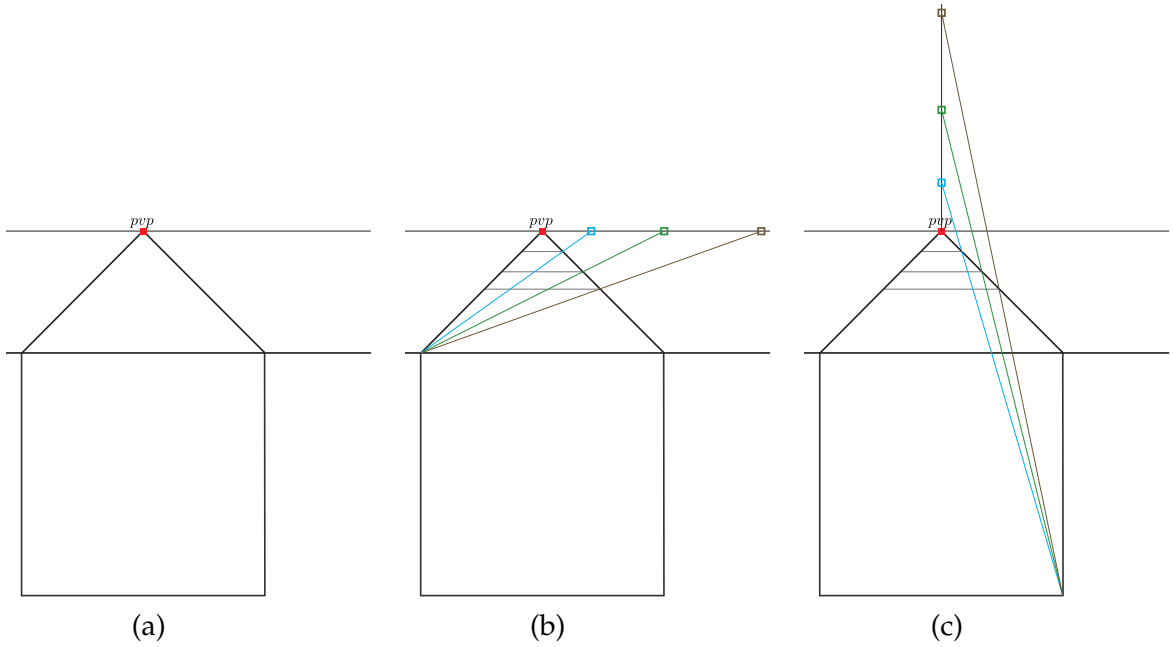


Figure C.5: (a) the viewer facing location sets the perspective tiled floor orthogonal edges. (b) different dps, i.e., viewer distances terminate the square tiled floor at different height on the perspective picture. (c) alternatively the corresponding auxiliary eye point can be used to determine depth limits of the tiled floor observed from different distances.

In practice, it is convenient to place the dp on the side of the picture. The dp is principally where the diagonals of the square tiles in perspective converges at infinity. Knowing the dp's position on the horizon determines the view configuration of a perspective picture. It is a convenient and intuitive rotated eye point, as demonstrated in Chapter 4, but it is not a center of projection between the perspective tiled floor and the physical tiled floor, as it is not invariant as the baseline is between the two half spaces. In the next section, I briefly show that in the formalism of Chapter 6, indeed has the auxiliary eye point invariant.

C.2 Auxiliary Eye Point Invariance

In Chapter 6 the following transformation was derived to map a perspective floor point, $\mathbf{Q}$, to its physical floor point representation on the canvas, $\mathbf{Q}'$:

$$\begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\frac{1}{\mathbf{E}_d} & 0 & 0 \\ 0 & \frac{1}{\mathbf{E}_h} & 0 \\ 0 & \frac{1}{\mathbf{E}_h \mathbf{E}_d} & -\frac{1}{\mathbf{E}_d} \end{pmatrix} \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \\ 1 \end{pmatrix}. \tag{C.1}$$

The auxiliary point is at $(0, \mathbf{E}_h + \mathbf{E}_d)$. Consider more generally the points on the central vertical axis passing through the pvp, that are points, $\mathbf{Q}$ of the form $(0, y)$. Using the above transformation we obtain,

$$\begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\frac{1}{\mathbf{E}_d} & 0 & 0 \\ 0 & \frac{1}{\mathbf{E}_h} & 0 \\ 0 & \frac{1}{\mathbf{E}_h \mathbf{E}_d} & -\frac{1}{\mathbf{E}_d} \end{pmatrix} \begin{pmatrix} 0 \\ y \\ 1 \end{pmatrix} \sim \begin{pmatrix} 0 \\ \frac{y}{\mathbf{E}_h} \\ \frac{y - \mathbf{E}_h}{\mathbf{E}_h \mathbf{E}_d} \end{pmatrix} \sim \begin{pmatrix} 0 \\ \frac{y \mathbf{E}_d}{y - \mathbf{E}_h} \\ 1 \end{pmatrix}. \tag{C.2}$$

Using $y = \mathbf{E}_h + \mathbf{E}_d$ we get the following physical correspondence:

$$\begin{pmatrix} \mathbf{Q}'_x \\ \mathbf{Q}'_y \\ 1 \end{pmatrix} \sim \begin{pmatrix} 0 \\ \frac{(\mathbf{E}_h + \mathbf{E}_d)\mathbf{E}_d}{(\mathbf{E}_h + \mathbf{E}_d) - \mathbf{E}_h} \\ 1 \end{pmatrix} \sim \begin{pmatrix} 0 \\ \mathbf{E}_d + \mathbf{E}_h \\ 1 \end{pmatrix}. \tag{C.3}$$

The auxiliary eye point, which is a special point above the horizon in the perspective $\mathbb{RCS}$, is indeed invariant in the transformation used to map between both floors. Indeed, all the points on the central axis above the horizon height, i.e., $y > \mathbf{E}_h$, are special. In effect, the eigenvalues of the matrix in Equation C.2, are the auxiliary eye point and the set of points on the baseline, the line spanned by the eigenvectors.

C.3 Double Projection

My framework includes projected shadows, which are a double projection problem. Artists have depicted them using construction lines on the canvas: Cole describes the properties in play in the geometry of projected shadows with examples on the ground plane, on steps, using different illuminations, including a sun on the side or a pseudo-light. Cole cannot exhaust all the configurations; it is expected that artists gain their independence by those studies and can adapt recurrent principles to any given configuration.

In my framework construction for Figure 5.25 (b), the shadow projection of a vertically rotated panel by angle θ onto a block medial face, I have specifically devise from the knowledge I built from Cole's examples. Thus, it is worth analyzing its result to validate the correctness of the construction. The geometry is intricate enough that is not obvious to see the construction is correct, thus revisiting it is necessary. Figure C.6 shows the pre-condition check: the rotated panel non-vertical edges converge to a single point on the horizon, vp_θ.

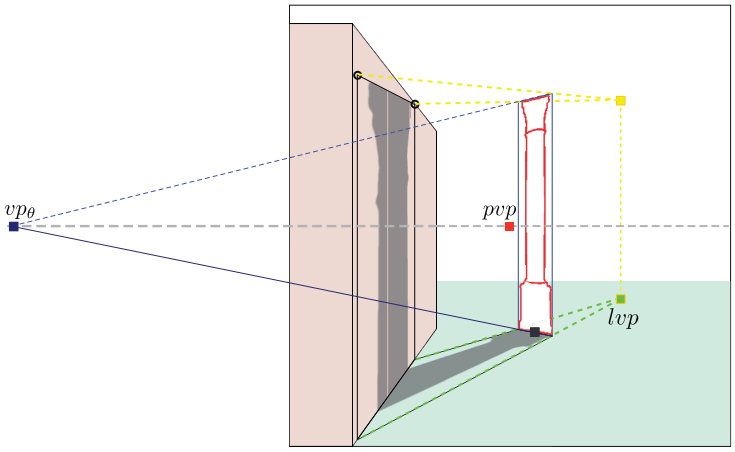


Figure C.6: Panel vertically rotated by θ projecting a shadow on the medial face of a block. vp_θ is the convergence point on the horizon of the lines rotated by θ.

C.3.1 Construction Line Analysis

Should the panel shadow projection on the block medial face have its boundary top edge converging at vp_θ? Obviously not. Figure C.7 highlights the geometry in play by extending construction lines. They indicate the following relations.

1. The top edge of the projected shadow on the medial face is extended to the horizon: the black line intersects on the other side of the pvp. The block medial plane points to $pvp = vp_{90^\circ}$. As the medial plane faces the panel rotated plane, which has the light behind it, a reflection applies.
2. The light rays to the panel boundary bottom corners on the ground intersect the plane of the medial face below the ground: on the medial plane the plain brown line is the projection of bottom line of the rotated panel boundary for the light at the center of projection.
3. The projected bottom edge of the rotated panel boundary on the block medial plane intersects the ground line to the pvp, at the same location than its projector on the ground extended to vp_θ: the red square.
4. The plane of the rotated panel meets the plane of the block medial face on the vertical of the invariant point (the red square of relation 3): the red line segment.
5. On the medial face the top edge of the projected shadow boundary (the black line of relation 1) meets at the end-point of two plane intersection segment (red line segment of relation 4): the red circle.

Those construction line relations are sufficient to convince one that the projection geometry used has consistency. Yet the description does not reveals all the dependence. Examining further, Figure C.8 (a) shows that the horizon intersection of the projected shadow bottom edge on the block medial face is connected to the light vanishing point.

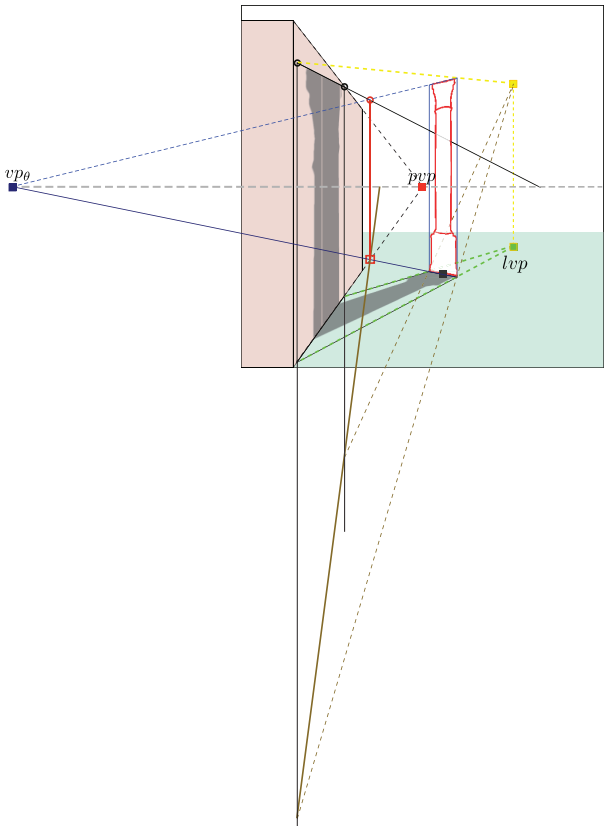
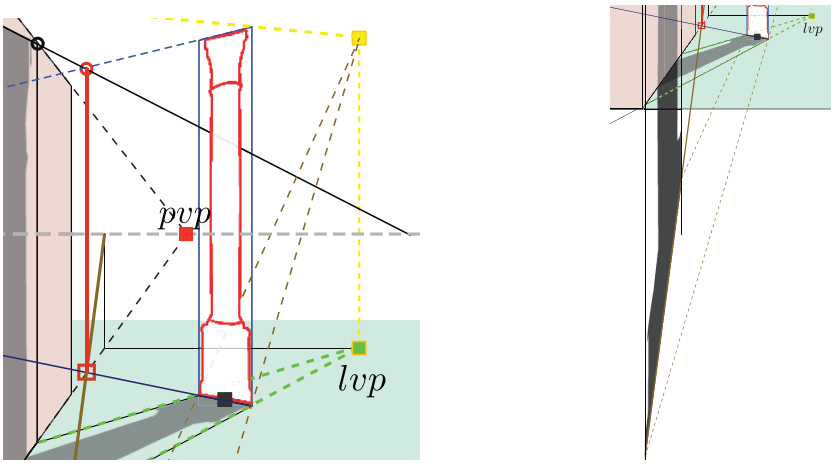


Figure C.7: Construction lines showing the relation of the geometry with the vertical segment that is the intersection between the panel rotated plane and the block medial plane.



(a) (b)

Figure C.8: (a) shows that the panel bottom edge projected on the medial plane is related to the lvp: its horizon point (indicated by the brown line) is vertically aligned with the lvp depth on the medial plane. (b) shows the unprocessed shadow below the ground.

C.3.2 Desargues Connection

From the analysis of the construction lines, we are now ready to examine the geometry in terms of Desargues theorem. The question is: where are the center and axis of perspectivity that define the projection of the rotated panel onto block medial face plane?

Figure C.9 shows two triangles in Desargues' theorem configuration, each of which are on the planes of the projection examined: $\triangle ABC$ is on the panel; $\triangle A'B'C$ is on the medial face plane. The respective vertices are related by being on the same light rays: the center of perspectivity is at the light source (the canvas location V on Figure C.9). The intersection L of AB and $A'B'$, M of BC and $B'C'$ and N of AC and $A'C'$ are on a same line: the axis of perspectivity is where the two planes intersect (the red line on Figure C.9).

In effect, it seems that the lvp determines in which vertical area of the medial plane the shadow projects. The light source on the canvas, V, determines where on that facing plane the shadow starts below the ground and where it ends. The orientation of the non-vertical edges, seems to be an effect of both. There is much going on in this double projection that would need more exploration between the 3D configuration and their relating canvas geometry to entirely grasp of the effect. However, I find that having construction lines connecting the shadowing plane and it shadow in my framework helps the control and the predication of the shadow effect, when the illumination points are moved continuously on the perspective picture.

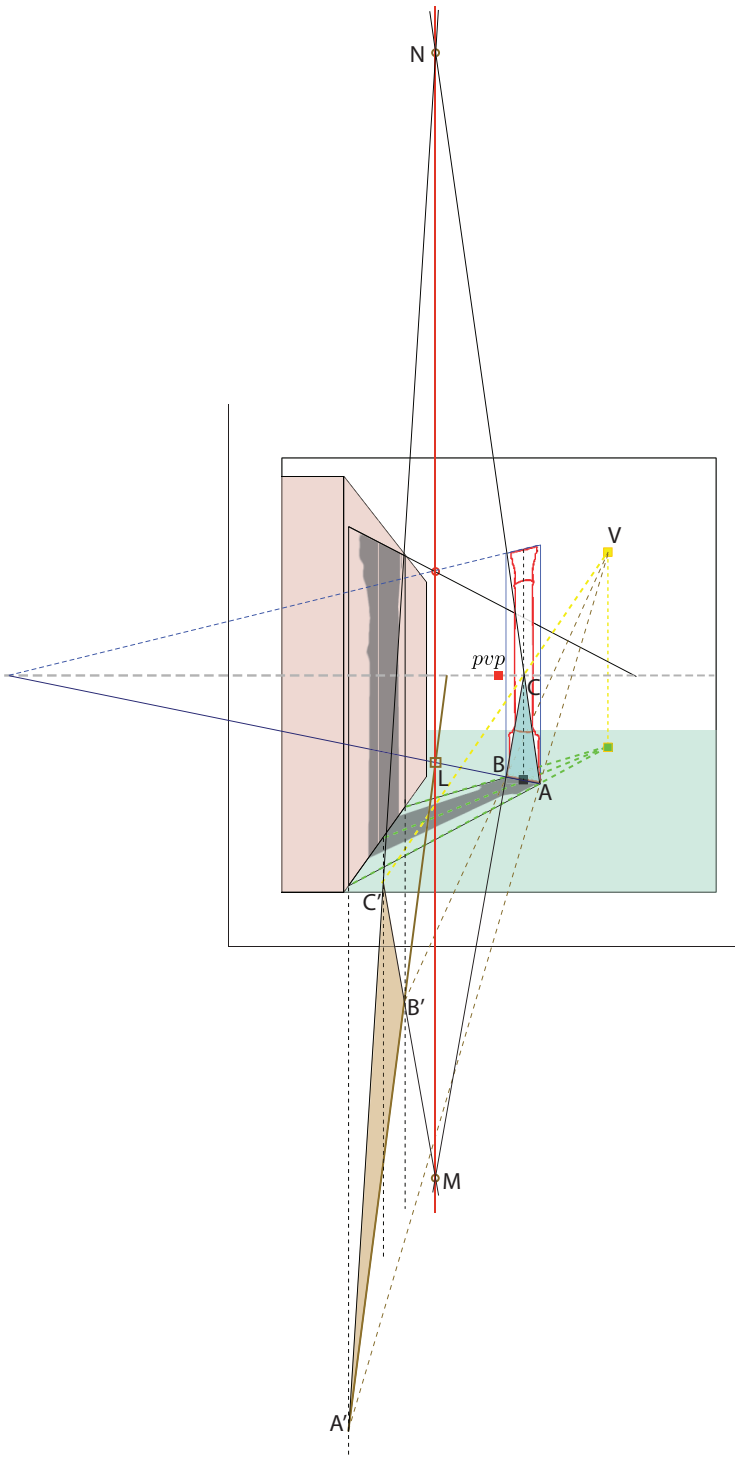


Figure C.9: Two triangles in Desargues' theorem configuration between the two planes of construction. Their center of perspectivity is V; their axis is the intersection line between the two planes.

Appendix D

Implementation

This appendix complements Chapter 5, by describing the user interface, data structures and rendering used by my framework.

D.1 Application Window

The graphical user interface is a window frame with a menubar, the content pane of which is divided into three parts, as shown in Figure 5.1. The column on the left side contains panel content, the images created off-line that are local representations of foreground objects. The main central region is the canvas, where the schematic representation is created. Below are a series of tab-panes grouping controls affecting different aspects of the picture.

While direct manipulation on the perspective picture is privileged, selecting by mouse interactions the perspective and lighting parameters, the position and size of volumes and panels, some actions require dedicated widgets. The tab-panes control some appearance and computation, plus interaction modes. Each tab groups the controls for one type of element of the picture: the planar pattern, the tiled floor, the objects (volumes and panels), shadows; plus overall controls (such as display and save).

The window frame has a preferred size and a minimum one from which it can be extended. The canvas occupies an important fraction of the total size; the picture is recomputed in function of the planar pattern limiting dimension, either width or height, that fits in the updated canvas size. The canvas content can be saved as a pdf file, in which form most figures in this thesis have been stored, using the SVG functionality of the Apache™ Batik Project [10]. The

floor lines, volume face paths and filled-in regions are of vector types, while the panel content and their shadows are of raster format.

D.2 Composition and Perspective

The planar pattern is sized to fit maximally inside the canvas area. Within it the perspective tiled floor is specified by manipulating a few parameters. There are global controls to translate and zoom the picture.

D.2.1 Planar Pattern

On the planar pattern tab-pane the user specifies the combination of vertical and horizontal basic pattern for the picture frame: with radio buttons ranging from one to three. The pattern and its recursive sub-patterns can be turned on and off to provide guidance for the 2D composition. Lists of snap points for each sub-pattern are kept, there are the locations of pattern line intersections. They form snap points sets that help the user to place elements accurately in geometric relationship for the picture 2D composition: Figure D.1 shows one such set to which the panel center has been repositioned accordingly.

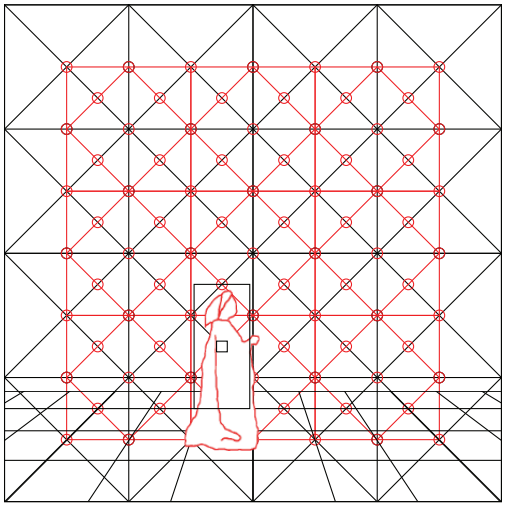


Figure D.1: The snap points of the first level sub-pattern of the planar pattern are displayed. The panel has automatically been moved for its center to coincide with the sub-pattern intersection that was closest beforehand.

D.2.2 Principal Perspective

Subsequent to selecting a planar pattern configuration, the user specifies the following within the constrained nature of the tiled floor in perspective: its discretization, its foreshortened appearance and its position in the planar pattern. The tiled floor tab-pane provides interface widgets controlling some of the floor parameters, which are used in conjunction to the principal perspective points defining its view appearance, directly available within the canvas area.

A slider adjusts the discretization of the tiled floor along the picture width: the number of tiles dividing the baseline into equal units. Another slider adjusts the tiled floor depth: the number of rows of tiles. The foreshortened appearance of the floor within the planar pattern is tuned by selecting from a variety of points the location of the distance point with respect to the principal vanishing point and the location of one of the vertical edges of the picture frame. Due to the central perspective picture symmetry, those specifications on one side of the *pvp* are sufficient. The framework displays potential locations for the distance point and for the picture corner on the right side, as shown in Figure D.2: unfilled rectangles along the horizon and along the baseline, the user selects each one by a mouse click within one of the rectangles of the set. Those discrete positions assure that the planar pattern is related to the tiled floor by some small integer ratio: the pattern diagonal to the frame corner, often a planar pattern diagonal, intersects the horizontals at subsequent heights which form a geometric series.

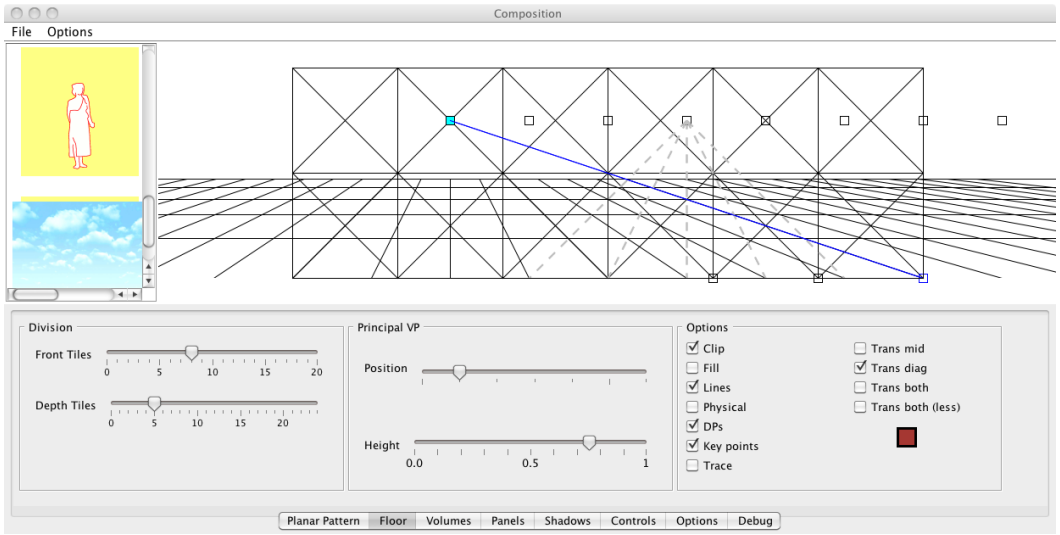
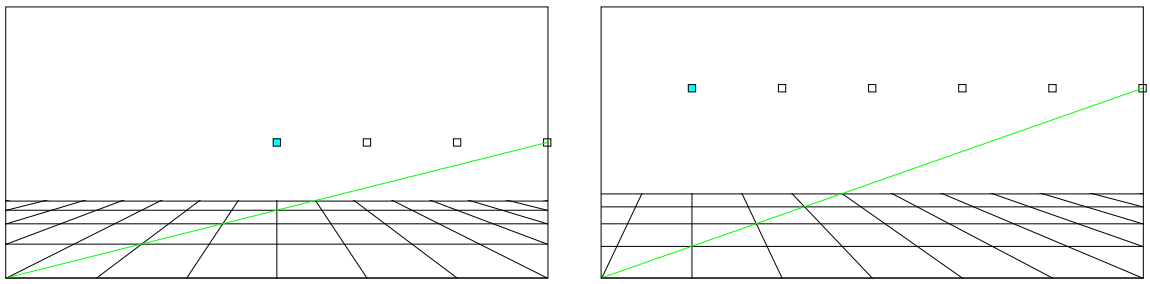


Figure D.2: Tiled floor tab-pane controls: the sliders defining the tiles discretization and *pvp* position. The picture is not cropped by the picture frame.

Another pair of sliders allow the user to move the principal vanishing point away from the usual center: the horizon height can be changed and the *pvp* position can be moved by discrete steps of the tile size. Figure D.3 (a) shows the tiled floor in the default centered view within the picture frame, which has been modified in Figure D.3 (b) to an asymmetric appearance within the picture frame. Those options allow reducing the stability of the picture central perspective: the distinct geometric center of the picture and the perspective convergence center diminishes the dominance of the picture centric system [6].



(a) (b)

Figure D.3: A tiled floor with six tiles on the baseline divided and four in the scene depth. In (a), the *pvp* is at the picture geometric center, the *dp* selected on the frame vertical edge. In (b), the horizon is raised above the picture geometric central axis and the *pvp* is moved by two tiles to the left, the *dp* remains on the right vertical edge. In (b) the tiled floor is then observed from higher and farther.

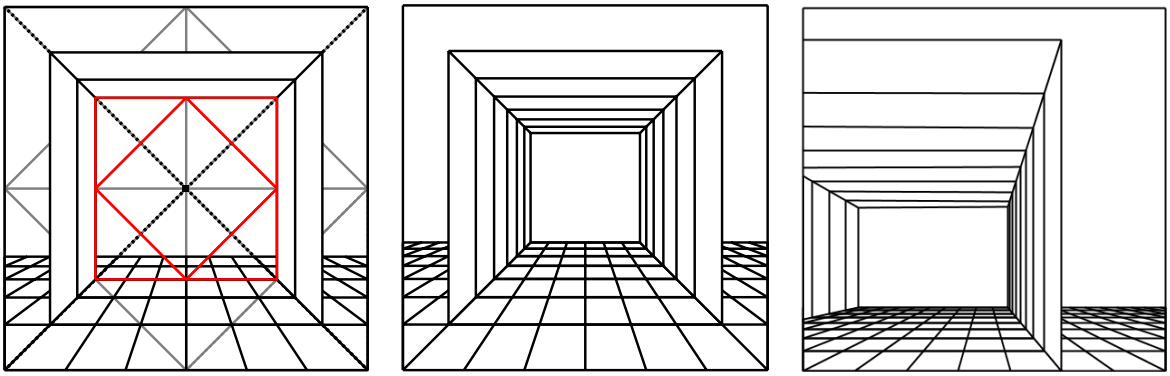
Finally, the framework also provides snap points relative to the tiles geometry. These snap points are useful to place exactly attachment points or origins in geometric relationships that have significance with the depth perspective of the picture.

D.2.3 Walls

A simple feature that links the picture perspective simulating 3D realism and the picture planar pattern governing its 2D composition is a wall feature based on Talbot's description [102]. An opened room volume encloses the overall space for a particular picture (it is used in Figures 8.2 and 8.5). The room geometry is constructed with its facing wall in reference to Talbot's planar pattern, its side walls in accord to the tiled floor appearance.

The back wall is defined according to a temporary wall, made of a subset of the first level inscribed squares of the planar pattern, which base intersects the tiled floor. The user defines the temporary wall dimensions selecting a sub-pattern rectangular region. Side walls are then

created, based on rays from the pvp and passing through the top corners of the temporary back wall, the dashed lines in Figure D.4 (a). Vertical division of the lateral walls are drawn with respect to the tiled floor horizontals. Sliders control the in and out range from the intermediate back wall for the length of the side wall, as shown in Figure D.4 (b). This room volume in perspective adheres to the tiled floor horizontals and to the planar pattern: it does not explicitly uses the dp.



(a) (b) (c)

Figure D.4: (a) shows the intermediate back wall sets for (b) the extended room volume where floor horizontals before and beyond for its lateral side complete it. (c) shows a semi-opened room volume that is not centered.

D.3 Objects

The geometric constructions that depict and manipulate blocks and panels that make up the perspective picture elements have been presented in Chapter 5 (§ 5.4.1 and 5.4.2), the interface and underlying data structures used are now reviewed.

D.3.1 Panels

This section describes the interface for the user to manipulate the panels' position and appearance in the picture. The data structures supporting these operations are also reviewed.

Interface. The user includes a panel in the composition by a mouse drag-and-drop gesture, moving a panel content from the left column of the application window to the canvas of the

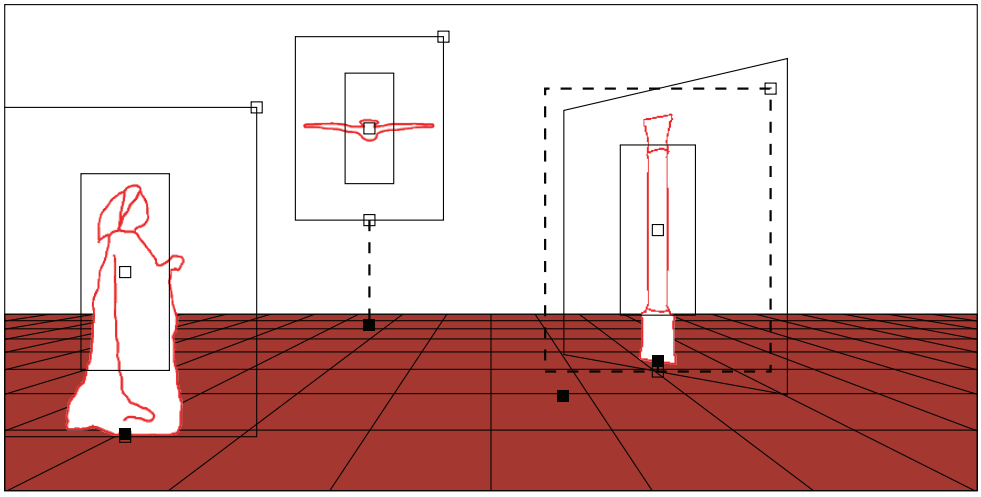


Figure D.5: Three panels with their manipulation handles visible. The column is rotated by 45°: its default frontal orientation boundary is the dashed rectangle; the unconnected small solid rectangle close to the second horizontal indicates the location of the bottom left corner of the image containing the rotated content.

perspective picture. When inserting such content the panel geometry is created: manipulation handles permit changing their appearances. They are the panel attachment point, origin, center, and scale rectangles, plus a selection rectangular region; all which provide direct affordances within and around the panel area.

By default, panels are parallel to the image plane and the user determines for each an explicit location within the perspective picture by specifying it laterally (its origin) and in depth (its attachment point): the attachment point defines the depth separating its vertical mid-axis from the baseline and the origin in relation to it separating the panel base starting height from the tiled floor.

As shown in Figure D.5, to reduce conflicts while picking a panel among other shapes, a limited region defines the selection rectangle: clicking within selects the panel. A selected panel is moved or scaled by direct manipulation in either 2D or 3D mode, as described in § 5.7: one works in the frontal plane, while the other functions in the picture depth. The scale handle is on the right corner of the panel's boundary rectangle. The panel center can be moved within the selection rectangle: the center defines the location to snap in 2D to the planar pattern intersections; in 3D the panel origin is snapped.

Finally, a panel can be vertically rotated to simulate a change in orientation of the object volume with respect to the viewer's line of sight using a slider for the angle. This effect is

achieved by using the 2D projective transformation that rotates the panel boundary in the perspective picture space. Knowing the positions of the four corner points of the panel from the default parallel orientation to its vertically rotated boundary by θ (Figure 5.16 or§ A.3.3), is sufficient to define the 2D projective mapping that warps the panel content [52]: a 3×3 projective matrix uniquely defines the transformation. It is applied to filter a copy of the panel content so to fit the quadrilateral panel boundary. As the filtered content is an image of rectangular dimensions, with transparent pixels outside the quadrilateral boundary, it is drawn according to the location of its extreme left corner (as the unconnected square belonging to the column rotated panel in Figure D.5 indicates). This transformation was not available to Renaissance artists. They transformed a grid drawn on the panel and used it as a guide for hand drawing, as they often did to draw anamorphic distortion [65, p.210]. On a computer the process is easy because it is linear and corresponds conceptually and mathematically to the operation of texture mapping hardware.

Data structures. The framework keeps a list of the panels included in the picture. For each user interface action, the implementation determines if a panel is affected and changes its attributes as requested: moving it in space, changing its size or orientation, or deleting it for examples.

In addition, each floor tile maintains a link to every panel whose attachment point is inside its boundary. Implementation-wise the inside tile test is easiest to compute on the physical representation of the tiled floor using the physical $\mathbb{RCS}$ attachment point. The tiled floor references to panels are fundamental to paint the object with proper occlusion among the other scene objects while rendering the scene using a planar painting algorithm.

D.3.2 Blocks

The block shape is the main way of implementing geometric volumes in my schematic representations for two reasons. First, the rectilinear shape approximates well general architectural forms. Second, the block shape underlies the rendering of the picture by splitting the scene depth space. Blocks are placed in hierarchical formations.

In this section, I propose an interface for block creation, describe the tree data structure that stores a block hierarchy suitable for approximating architecture, and explain the units of blocks, which are needed to render in the 2D context. Chapter 9 proposes for future work local adjustments required to depict other geometric volumes.

Interface. By default the orthogonal projection onto the floor of any block matches a rectangular set of tiles. Therefore, blocks that sit directly on the floor are created by elevating the set of tiles corresponding to its base by its height (§ 5.4.1). The user specifies the set of floor tiles in the block's base and the block's height, as it appears in the perspective picture.

To select the block base on the floor, mouse interaction is convenient: a mouse drag with the left button down for example selects a rectangular set of floor tiles—the intersection between the mouse cursor path and the tiles' interiors defines the selected base—the two extremes coordinates[1] of the path keep the selected tiles region rectangular. From the floor selected base the shape is creating using further affordances. The block height can be controlled by mouse wheel scrolls or alternatively by widget buttons in the control tab-pane.

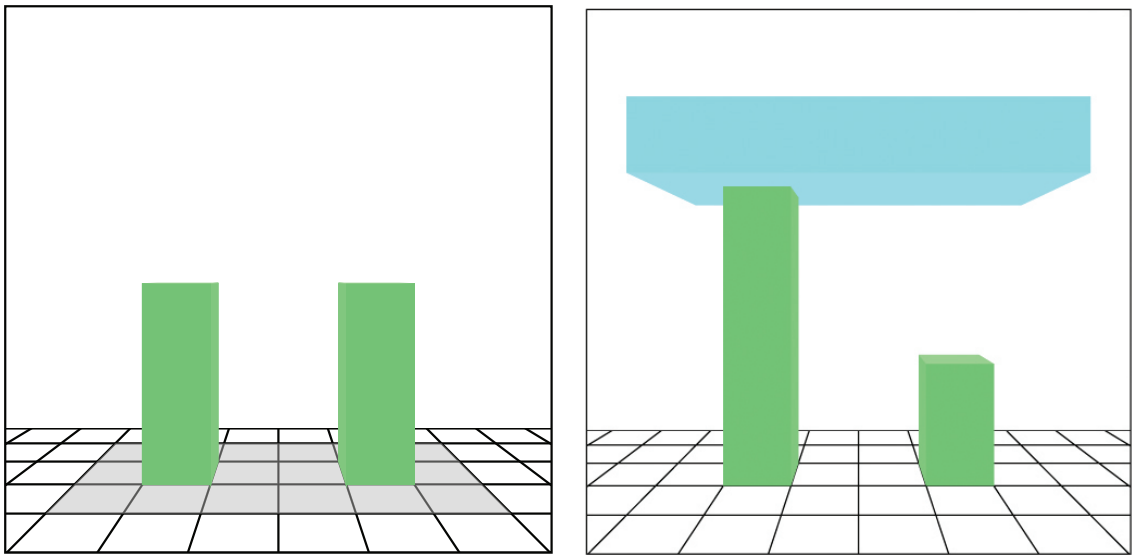
Hierarchical Structure. The approximation of scene architecture is better handled by a collection of blocks resting on top of each other: a lintel or architrave on a pair of columns requires it as well as a structure replicated atop so as to form a facade with several floor levels. Much architectural structure logically decomposes into parts, each a volume that is reasonably approximated by its bounding box.

Thus, the framework maintains a hierarchy of structures composed of blocks as a tree that is modified by user input. The blocks are the nodes of the tree and the tiled floor is its root. The simple blocks resting on the floor described in Chapter 5 (§ 5.4.1) are the first level of the tree and have the floor as their parent. At that level, the maximum number of children is the number of tiles. In practice, the floor has as many first level children as they are non-overlapping blocks resting directly on sets of floor tiles.

Having mentally decomposed into blocks an architectural geometry, the user then sequentially defines them in the framework from the floor up. The framework then constructs each block in relation to existing ones. Specifically, if a previously defined block overlaps the floor base of a new block, the framework constructs the new one atop of the previously constructed one: a specified base on the floor that intersects other base(s) implies the new block being above the highest reaching block, as if the new block was dropped from the sky landing on the first horizontal surface it touches. (If there are several previous blocks of the same maximum height, any of them may be chosen as parent.) A block has in effect two bases: the one projected onto the tiled floor and its actual base, parallel to it, but possibly differing from it in height (the total height of its parents to the floor).

Figure D.6 (a) illustrates the base selection that creates a block atop former ones: the floor

[1]Some included tiles have not been intersected by the mouse path.



(a) (b)

Figure D.6: Block hierarchical formation. (a) tiles specifying a new block are selected. (b) the new block atop after one of the block below has had its height increased.

base intersects the ones of existing blocks. Figure D.6 (b) shows the created block redrawn after one of the blocks below it is increased in height. The framework maintains the tree hierarchy as the user alters the parameters that adjust the piled structures the blocks form.

Pasting. Panel content can be used to texture block faces as is used in Figures 8.8. The content of a panel is moved to decorate plain surfaces, as shown in Figure D.7: the texture mapping is a simple way to add windows, since the surface details are negligible.

Repeating patterns are common. To ease their creation, block faces are split in regular physical parts. The panel content then repeat to create the overall appearance of the block face, as shown in Figure D.7. After selecting a block, the user picks one of its apparent faces (medial, front or top), and then a selected panel within the picture pastes its content appropriately according to the split number indicated. When a panel content is pasted on a volume face, the panel ownership is transferred to the volume structure. The volume is then responsible for rendering it as part of drawing the face.

Alternatively, a panel can be used as a back-drop for the scene . In that situation, its infinite depth of the special panel means it is drawn behind every object in the scene, i.e., it is drawn before anything else.

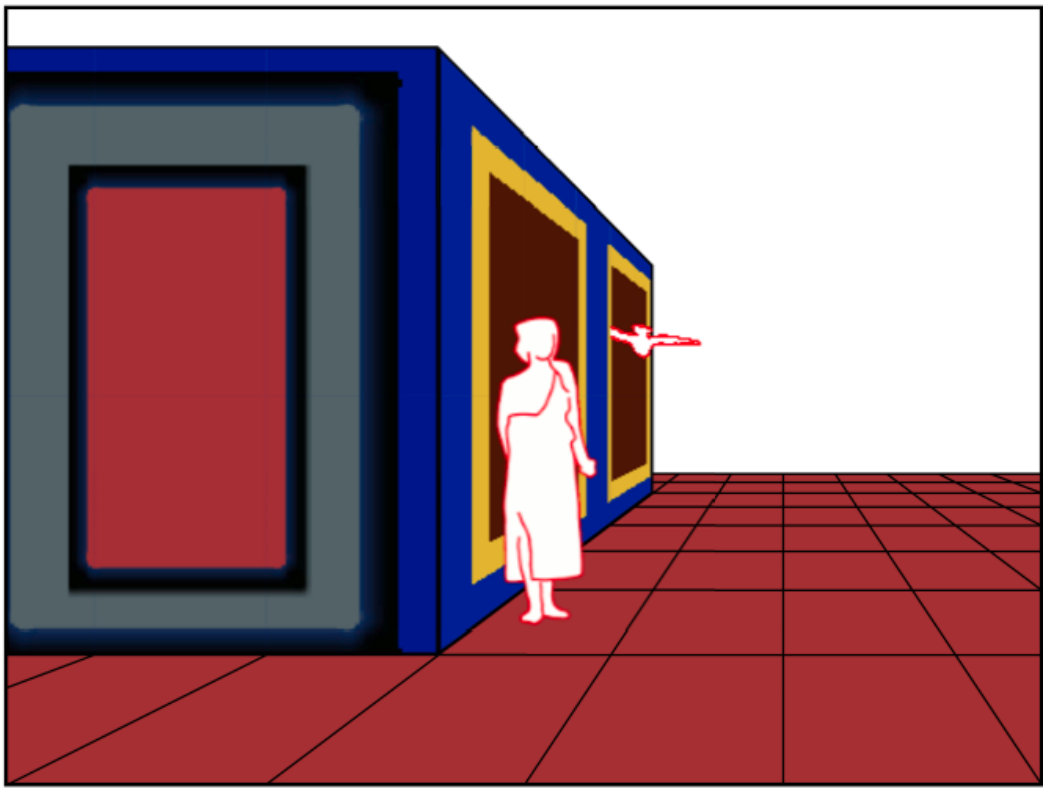


Figure D.7: Textures to decorate the faces of a block. The panel content can be repeated, it is then fitted within equal division of the face that is foreshortened.

Unit Block. General blocks have floor bases possibly made of multiple tiles, always a rectangular set. Unit blocks have a fixed physical base area, each defined by its projected base matching a single floor tile. Every general block is identical to an integral number of unit blocks. The unit blocks that comprise a block are the critical structure that permits rendering the perspective picture with correct occlusion (Chapter 7): in Figure D.7 the bird coming out of the window has the effect of its left wing partially occluded possible because of the rendering using unit blocks.

Consequently, additional data structures relate the blocks to the floor discretization. The floor tiling kept as a 2D array, each tile array element maintaining a list of links to the unit blocks whose projected bases equals the tile. The explicit rendering algorithm, Chapter 7, iterates the lists to draw the unit blocks making up the architecture volumes portrayed on the picture.

D.4 Projected Shadows

Projected shadows are part of the framework to study another projection relation that artists realize directly on the picture. Simulating them requires isolating the key canvas points that encode the projective relations in space among the planar surfaces that cast and receive the shadow and the source of illumination. The portrait of shadows help to reinforce the depth location of elements contained in the picture.

Considering this goal, it is sufficient to use a single source of illumination so as to not overcrowd the picture with many secondary shapes. Illumination locations and types are also restricted to those recurrent in artists practices because of the appealing effects they produce.

D.4.1 Masks

Projected shadows are simulated using mask images, which are automatically created from a semi-transparent panel content; the panel content filtered to a black and white image with an alpha channel: coloured pixels that are opaque are converted to white opaque pixels whereas transparent pixels are converted to black pixels. The mask is used to paint the shadow, the white pixels defining where the panel content blocks the light and the shadow should be painted; the black pixels are transparent when the shadow is rendered.

To produce a specific shadow, the mask image is transformed according to the geometry that projects the panel boundary corners with respect to the light centers on the canvas (§ 5.5 and A.4) using the appropriated collineation, a 2D projective transformation. Given a panel shadow quadrilateral for a panel rectangular boundary, the mask image is warped using the unique 2D transformation relating them [52]. The transformed mask is a rectangular image: its width is the length between the minimum and maximum x-coordinates of the four projected corners its height the equivalent for the y-coordinates. The transformed mask image specifying the shadow interior needs to be drawn from the proper, possibly offset, bottom left corner location in the picture geometry: the shadow painted through the mask needs to relate (connected or not) to the shadowing shape. This 2D texture mapping mechanism automates artists transfer of a complex shape into its shadow using a deformed grid.

D.4.2 Appearance

There are a few parameters for modifying the appearance of shadows: shadow colour, opacity and sharpness of their boundary edge are available controls in my framework. Beyond the

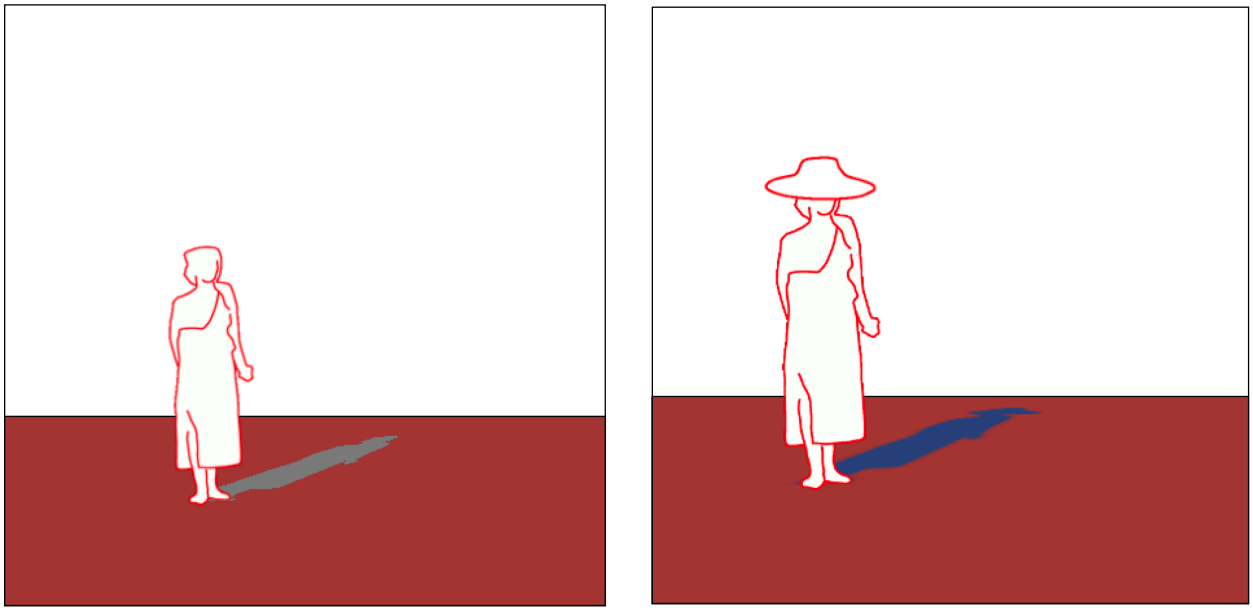
shadow geometry handled by the framework computations, the user can stylize their appearance so as to portray different illuminations. The shadow's parameters described below are used in the implementation in accordance with artists' mind set.

- Artists use the colour of shadows to reinforce the colour of atmosphere and surrounding contrasting surfaces. Rather than computing the inter-reflections, artists for example use directly bluish colours to paint their shadows so as to strengthen the effect of the overall sky illumination.
- Different illuminations create different shadow opacity. In a bright illuminated interior, lights bounce back from different surfaces: illumination is not direct and thus shadows from lamps are often transparent. The opacity of two shadows adds together when it comes from two sources, while it does not when it is the effect from the same source. In contrast, a street lamp post at night creates opaque shadows.
- Similarly the boundary of shadows differ from the environment condition. In some cases, shadows are closed by sharp edges, whereas in other situations shadows have fuzzy boundaries, making it hard to determine where they stop: a light torch creates the former, a fire place the latter.

On the shadow tab-pane, widgets are available to change the shadow's overall appearance: a colour swatch to change the shadow colour; each of the opacity and edge blurriness is controlled with a slider. When the opacity is changed, the shadow mask images are interactively redrawn with a different alpha setting. For the blurriness option, further shadow mask images are recomputed, using adapted matrix kernels that filter the non-transparent region of the panel content so as to produce fuzzy boundary edges. Figure D.8 illustrates the use of these simple parameters to add expression to projected shadows on the ground plane.

D.4.3 Rendering

Each panel in addition to having a reference to their image content, also has a reference to their masks. According to the shadow geometry computations, i.e., the light, among the objects making the picture, panels and blocks cast shadows on each other's surfaces. In the case of a shadowing panel, the shadowed block has a reference to the distorted mask image, so as to render it atop of its shadowed face. Blocks create simple projected shadows on the floor, the projection of each face, a distorted quadrilateral, is sufficient.



(a) (b)

Figure D.8: In (a) the default appearance of a projected shadow on the ground plane: the shadow geometry is handled by the framework; the user with the panel attachment point assures that there is a contact point. In (b) another panel is inserted to put a hat on the figure, the shadow parameters are used to beautify the illumination effect: a shadow is blue and semi-transparent, the edges are now smooth.

In addition, for semi-transparent shadows alpha blending needs to be handled so as the painting through the mask images do not add, as the shadows are simulated for a common light: composite images are useful to keep the computations efficient. Lastly, as discussed in the next section clipping is required to paint shadows on block faces.

D.5 Rendering Algorithm

This section describes the rendering algorithm implemented in my framework. Compared to Chapter 7 it includes the panels and textures that are part of my schematic representation. A pasted texture or a projected shadow on blocks on a block face, generally spans across faces of unit blocks. Panels are rendered according to their attachment points, each of which is within a tile.

D.5.1 Assumptions

My framework rendering works under the following assumptions.

- Overlaps exist in the perspective picture due to occlusion among scene objects, i.e., blocks and panels may occlude each other.
- Entanglements are not considered, i.e., no cyclic relations between scene elements are permitted. Blocks and panels are convex and separate entities in 3D space that do not inter-penetrate in a way that they are twisted in each other's volume.

Each object pair spatial relationship is characterized as behind, in front, on the side, on top and below.

D.5.2 Unit Blocks & Clipping

The rendering described in Chapter 7 gave three natural orders that can be used. In my implementation the tiles are processed row by row from deepest to closest, from the most dorsal tile to most medial one on each row. Iterating through the tiles, unit blocks on them are considered. There are two ways to consider each unit block.

One way is to use the unit block to directly paint: the actual blocks the user defined are thus simulated by painting the sets of unit blocks in a valid occlusion order. The first approach is useful when block faces are all plain: the drawing simpler and the performance better.

The second way is to use each unit block as a clip volume through which to draw the actual block: the actual block is drawn multiples time, each time using a different unit block for the clip, painting the faces which span tile boundaries progressively.

This second way is necessary when blocks' faces are not all plain: a texture or projected shadow is present on one face. A block face spans tile boundaries: the textures require splitting for proper visibility. Drawing the blocks using unit block clips in effect splits the textures automatically. This method is advantageous compared to decomposing each texture projection in sub-images with respect to tile boundaries. Clipping computations are discrete and independent of the scene occlusion complexity, and are handled in hardware by the JavaAPI. However, those clippings sometimes produce small artifacts that can be noticed (the thin line across the shadow on the medial block face in Figure 5.25 for example).

D.5.3 Algorithm

The steps below are the sequence followed by my framework rendering. The sequence has three parts: pre-floor, perspective picture, i.e., its objects on and above the floor, and post-floor. There are some optional settings depending on the display controls of the user which are omitted in the description.

Pre-floor. Setting the frame area and painting the far and ground planes.

1. The clip region is initialized to the frame boundary.
2. The back-drop is painted.
3. The clip region is set to the ground plane area: from the baseline to the tiled floor last horizontal.
4. Each of the following is painted.
    - The ground plane filled.
    - The ground plane projected shadows, which are all gathered on a composite semi-transparent image.
    - The tiled floor lines.
5. The clip region is re-initialized to the picture frame.

Objects. The objects are rendered sequentially processing the tiles on either side of the *pvp* (Figure 7.13): from the deepest row, most dorsal tile, inter-leaved by the reciprocate tile on the other quadrant side, moving closer to the picture central vertical, switching between quadrant as long as possible (the *pvp* may not be at the vertical center). The last processed tile on each row is on the picture central vertical if a tile spanning it exists; this successive approach is repeated row by row until the baseline. For each tile the following is done.

6. Unit blocks are considered from the frame horizontal edges as they successively approach the horizon from either side (Figure 7.14). If there is a unit block crossing the horizon, it is handled last (its horizontal faces omitted). For each unit block either its faces are successively painted or used as clip, according to the sequence described in § 7.4.
7. The panels whose attachment points are on that tile are drawn in decreasing depth order of the attachment point coordinate, and from the side to the center in case of tie.

Post-floor. Perspective picture construction lines or key points are painted.

This description includes the essentials, omitting details such as the drawing of block edges or the slight change of colour applied to face according to the light position. These are obvious features added to my framework.

In rendering a single image most of the drawing is lines. The transformations described above and in Chapter 6 are used to find endpoints and 2D graphics hardware draws the lines, automatically handling resampling and anti-aliasing. The remainder of the drawing is texture-mapping, which accounts for most of the pixels drawn. Texture-mapping is accomplished by using matrices to determine the images of four unique no-collinear points, usually the vertices of a quadrilateral, from which a projective transformation is calculated. Using the transformation 2D graphics hardware then handles the required resampling.

In fact, performance is not an issue when rendering simple images. A user, however, interacting heavily with my framework can experience a loss of response. In such a case, the right solution is to lower the quality of the rendering while keeping it timely: fine detail is invisible to users when an image is changing.

Appendix E

Tiled Floor Construction Variations

Chapter 4, § 4.2 described two constructions to draw a tiled floor in perspective: Figure E.1 compares them; Piero's tiled floor diagram with the actual intersections evaluated in the physical space, is juxtaposed with the 3D side view inherent to the diagonal point construction.

In addition to these two methods, there are many more variants [65]: most are geometrically equivalent, varying in efficiency or practicality. This appendix presents two such constructions: one is a slight variant on the diagonal point construction, without the requirement of tile edges aligned to the picture central vertical; the other duplicates lines to augment precision.

E.1 Many Methods

The most difficult aspect of depicting a tiled floor is estimating the location in perspective of each floor horizontal. Many methods existed in the Renaissance for doing so, having in common the use of at least a point of convergence on the horizon, away from the *pvp*, through which one of the two sets of the floor diagonals passes. Some methods that produce a foreshortening effect in the placement of the floor horizontals are not equivalent.

In effect, separating the convergence point from the *pvp* by a length determined by the distance from the viewpoint to the canvas is not always followed. As such, this convergence point cannot always be called *dp*. Some instances of the construction use other successions of diminishing proportions, such as the intersections of the floor diagonals with one of the vertical edges of the canvas frame, when the convergence point is positioned outside the canvas boundaries, for example.

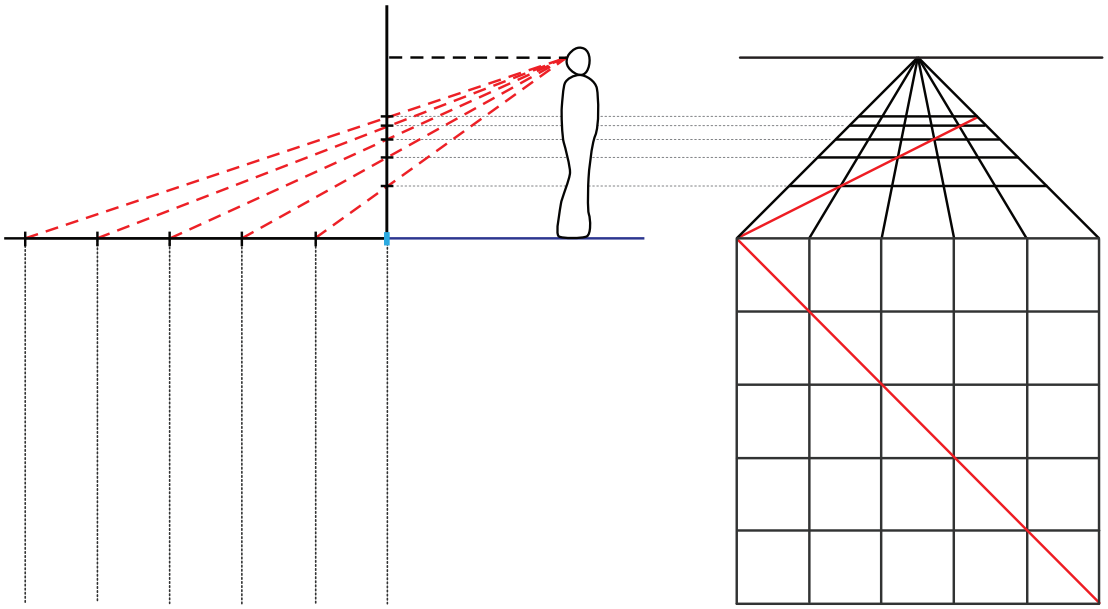


Figure E.1: To emphasize the 3D and 2D correspondence of the floor construction, two images are put in comparison: on the left the simple 3D side view of the viewer, the image plane, and the tiles in depth; on the right is the corresponding canvas image plane viewed frontally, as presented in Piero's diagonal construction diagram. The dashed lines demonstrate the correspondence of measures between the two split views.

The two alternate methods for constructing the tiled floor, described below, share the convergence point of the floor diagonals as the distance point, meaning that it represents the distance separating the viewer from the canvas, the same 3D to 2D correspondence of the two methods described in Section 4.2.

It is remarkable that the ingenuity of the distance point representation leads to a variety of equivalent constructions, demonstrating the robustness of its paradigm. Renaissance artist practices and writings assess that an awareness of the distance point power existed to some extend.

E.2 A Distance Point Variant

Piero's construction reveals that it is unnecessary to require a vertical floor orthogonal from the baseline to the *pvp*. As a matter of fact, the distance point construction described in Sec-

tion 4.2.2 generalizes to many equivalent methods using other floor orthogonals, since either one of them can be selected as the image plane reference.

For example, a convenient alternative uses the orthogonal closest to the distance point, since most baseline tile corners exist beyond it. Obliqueness or not, all floor orthogonals separate the canvas space into two equivalent halves. One half contains the eye, the other a subset of tile corners on the baseline, which are used to evaluate the successive depth levels of the tiled floor. Regardless of the floor orthogonal chosen, the intersections obtained are at the same height on the canvas, as shown in Figure E.2. Using a small section on one side of the canvas, rather than the middle or the entire canvas, for the construction has its advantage.

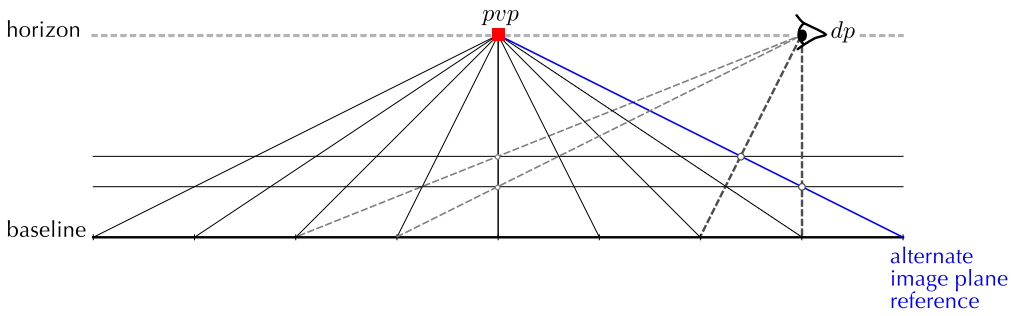


Figure E.2: Either one of floor orthogonal can be used as the image plane reference for the distance point method previously described. Equivalent intersections are obtained, as the same heights are evaluated since the floor horizontal are parallel including to the image plane.

Figure E.3 demonstrates the construction with a viewpoint close to the canvas, which creates an elongated floor. Piero's diagonal to the other side of the floor is drawn to show the agreement between the methods.

E.3 Pair of dp's Construction

Another variant of the construction uses both distance points, in combination with opposite baseline corners. The two dp are on the horizon, equally distant from the pvp, and on either side. With this construction, no floor orthogonal intersections are needed: intersections between the two sets of floor diagonals create sufficient landmarks. Figure E.4 is an illustration of this alternate method. It is worth noting that intersections at the center of tiles can be used to refine the floor further, with the potential to depict a floor of rectangular tiles, for example. In addition, the segment in between the two dp is the span of a 90° viewing angle on the

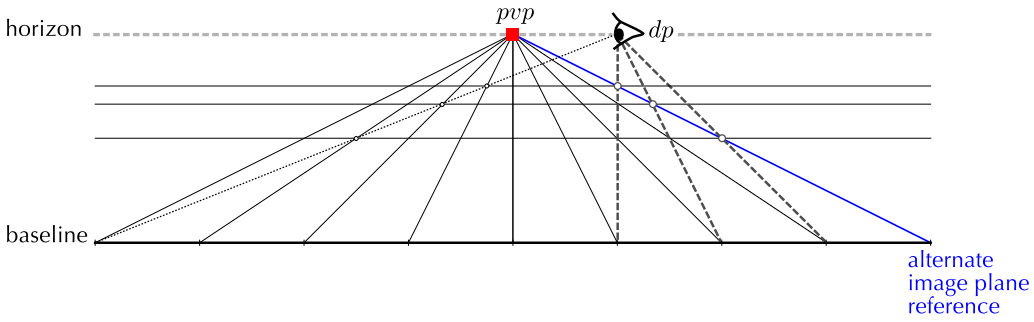


Figure E.3: When the distance point is close to the principal vanishing point, the floor diagonals joining the first few baseline tile corners, behind the closest floor orthogonal, might be behind the eye. Nevertheless, the dp's paradigm still holds.

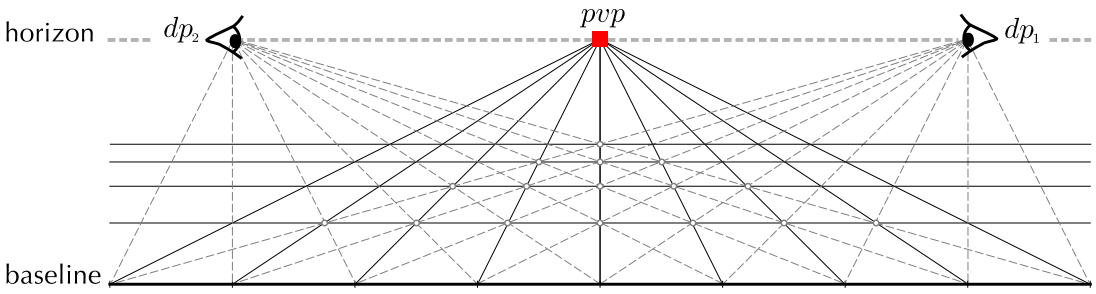


Figure E.4: Another flavor to constructing the tiled floor uses both distance points on either side of the principal vanishing point. The mutual intersections of their floor diagonals produce all the information needed to create diverse perspective floors.

canvas. Indeed, one set of diagonals are physical lines at 45°, clockwise with respect to the view direction, while the other set is also at 45°, but counter-clockwise. Each set converges to the appropriate dp on the canvas, so that the length between them on the horizon is the projection of a 90° field of view on the image plane.

E.4 Practical Implications of The Diverse Constructions

The two tiled floor constructions, described in Section 4.2.2, and the two variants, described in this appendix, demonstrate the great variety of equivalences available within the distance point paradigm. Those equivalent methods had practical implications for Renaissance artists.

Indeed, the different constructions do not compete with one another, but are complementary. They provide different levels of precision: using a single intersection, for example, compared to using many.

Piero's unique diagonal is a quick method to verify a tiled floor built with a different construction. Using multiple lines joining separate measures, creating many intersections, reduces the precision required for each placement of the straight edge. In fact, the two sided construction, Section E.3, is probably the most accurate, as multiple intersections for each depth level are created, averaging out possible errors. Furthermore, the intersections are closer to 90°, which make them precise. Thus, using a combination of constructions in a floor manual execution refines the accuracy achieved.

Renaissance artists faced with the practical problem of building a tiled floor, which is essential for developing further the perspective of the scene, created many constructions, differing only in the measurements needed and the construction lines drawn. Their practices show methods of using the distance point that evolved to allow different trade-offs between time spent and accuracy achieved.

Specifically, my review of different methods based on the distance point concept to construct the tiled floor, demonstrates important relationships between the spaces as

- in the scene versus on the picture, or equivalently as,
- in the 3D physical space, versus on the 2D perspective space.

Piero's juxtaposition of floors, which shows how to mingle both space on the flat canvas, is essential to simulate 3D with depth realistically constructed while working solely on the image plane: the algorithms of my framework, described in Chapters 5 and 6, are based on it.

Appendix F

Construction Of An Octagonal Volume

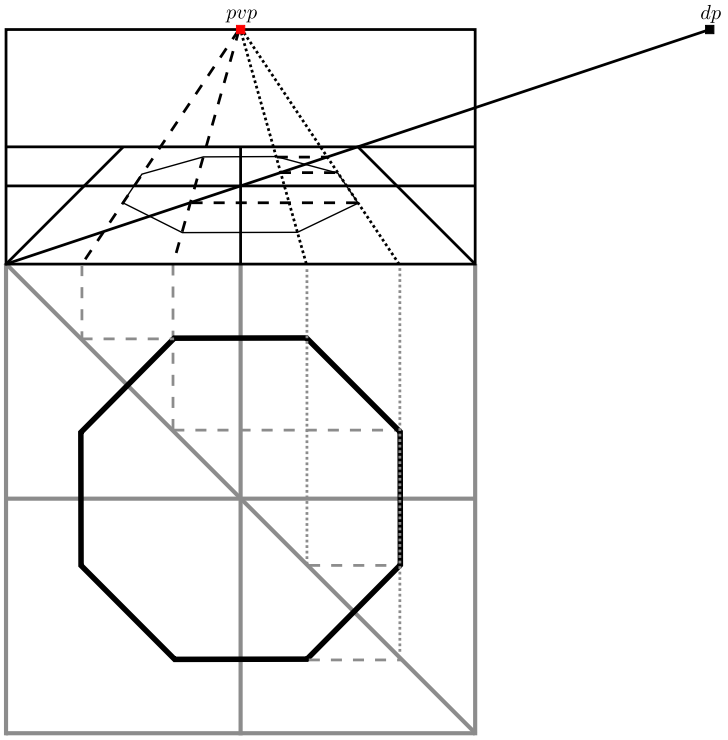


Figure F.1: The octagon drawn easily on the physical floor is transferred to the perspective space using the geometric construction of Section 5.3.

F.1 Octagonal Base

In Figure 4.12 (a), Piero showed how a rotated octagonal base is depicted in perspective: the octagonal shape is first drawn in the physical space; it is then mapped to the perspective space, vertex by vertex. The construction described in Section 5.3 provides the mapping between a physical arbitrary point to its perspective one.

Thus, the framework implements an algorithm for computing the perspective appearance of any geometric shapes on the ground. Figure F.1 shows it constructing an octagon aligned with the tiled floor lines. The same applies for any polygon defined on the physical ground.

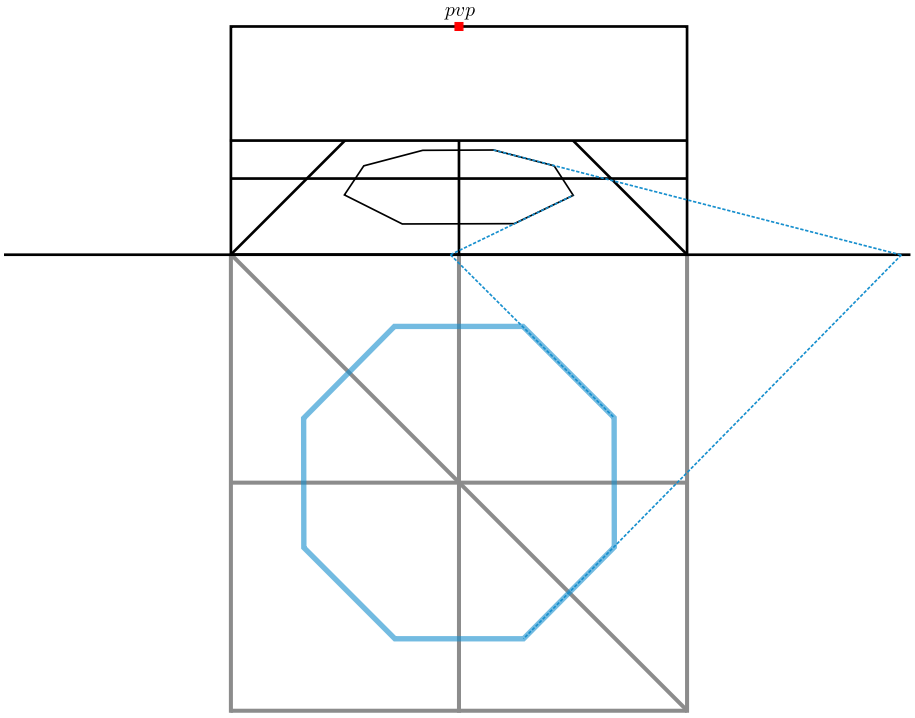


Figure F.2: A line on the perspective floor *is the image of* a line on the 3D physical floor: its representation. The correspondence between the two floors means that the perspective line and its 3D physical line on the floor, rotated on the image plane, intersects at the baseline.

While the construction involves for each vertex a set of floor lines joining at Piero's diagonal and at the baseline in each of the two floors, an efficient consistency check exists. In effect, each line that is not parallel to the baseline intersects the baseline at the same point as the line on physical floor it corresponds to does. Floor orthogonals, diagonals, and in fact for any line orientation that is not parallel to the baseline, their representations meet at the baseline. That

is because the baseline is a privileged location, shared by both spaces. Figure F.2 shows this baseline property for two corresponding edges of the octagon.

F.2 Octagonal Volume

Using a main orthogonal and its elevated version according to a given physical height, the foreshortened dimensions along the scene depth can be deduced, as shown in Section 5.4.1.

Thus, given the perspective base shape of a volume, as produced by the method of Section 5.3, any volume elevated from the base can be depicted in perspective. The only difference is that line intersections are required for each base vertex at a distinct depth to transfer them to the main orthogonal plane on the canvas side. On it, the foreshortened heights are evaluated.

Thus, with this elevated construction, the practice used by Piero to locate the vertices on an octagonal prism, as shown in Figure 4.12 (b), is automatized. Figure F.3 uses the perspective octagon base of Figure F.1 to elevate an octagonal prism, using such a method. The top face is above the horizon, and thus invisible.

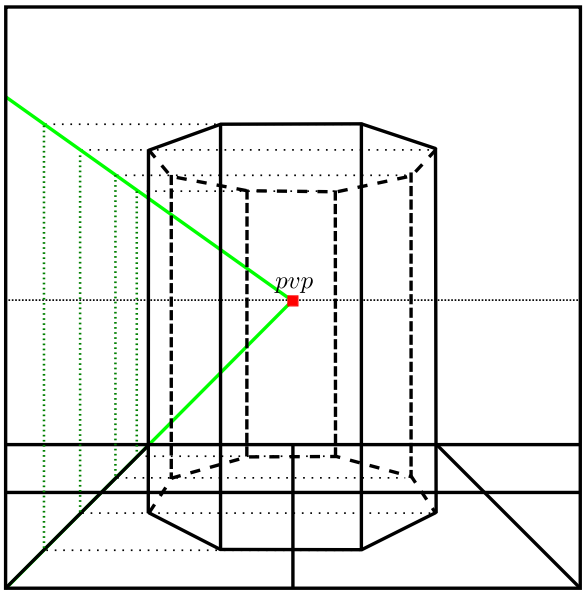


Figure F.3: From the octagon drawn on the perspective floor in Figure F.1, an octagonal volume is elevated using the construction of Section 5.4.1, evaluating each distinct foreshortened height on the main orthogonal plane.

Bibliography

[1] Adobe Systems, *Adobe Illustrator: Help and tutorials*. `http://helpx.adobe.com/pdf/illustrator_reference.pdf`, Illustrator CS5-CS6 (accessed December 2012). 28

[2] M. Agrawala, D. Zorin, and T. Munzner, *Artistic Multiprojection Rendering*, in Proceedings of Eurographics Workshop on Rendering Techniques, 2000, pp. 125–136. 25, 183

[3] L. B. Alberti, *De Pictura (On painting)*, 1435. 42, 55

[4] K. Andersen, *The Geometry of Art: The History of the Mathematical Theory of Perspective from Alberti to Monge*, Springer, 2007. 55, 58, 71, 72, 226, 229, 230

[5] R. Arnheim, *Art and Visual Perception: A Psychology of the Creative Eye*, University of California Press, 2nd ed., 1974. 51

[6] ——, *The Power of the Center: A Study of Composition in the Visual Arts (The New Version)*, University of California Press, Berkley, Los Angeles, London, 2009. 81, 174, 176, 178, 244

[7] K. W. Auvil, *Perspective Drawing*, Mayfield Publishing Company, 1990. 49, 71

[8] V. E. Barnett and P. H. Barnett, *The originality of Kandinsky's compositions*, The Visual-Computer, 5:4 (1989), pp. 203–213. 30

[9] R. Barzel, *Lighting Controls for Computer Cinematography*, Journal of Graphics Tools, 2:1 (1997), pp. 1–20. 23

[10] Batik, *SVG Toolkit*. `http://xmlgraphics.apache.org/batik/` (accessed December 2012). 241

[11] X. Cao, Y. Shen, M. Shah, and H. Foroosh, *Single view compositing with shadows*, The Visual Computer, 21:8-10 (2005), pp. 639–648. 23

[12] R. Carroll, A. Agarwala, and M. Agrawala, *Image Warps for Artistic Perspective Manipulation*, ACM Transactions on Graphics, 29:4 (2010), pp. 127:1–127:9. 22, 27

[13] P. Cavanagh, *The artist as neuroscientist*, Nature, 434 (2005), pp. 301–307. 98, 192

[14] P. Cavanagh and Y. G. Leclerc, *Shape From Shadows*, Journal of Experimental Psychology: Human Perception and Performance, 15:1 (1989), pp. 3–27. 72, 110

[15] K. Clark, *Piero della Francesca: A Complete Edition*, Oxford University Press, Oxford, 1981. 38

[16] J. M. Cohen, J. F. Hughes, and R. C. Zeleznik, *Harold: A World Made of Drawings*, in Proceedings of the 1st international symposium on Non-Photorealistic Animation and Rendering, 2000, pp. 83–90. 21

[17] R. V. Cole, *Perspective: The Practice & Theory of Perspective as Applied to Pictures, with a Section Dealing with Its Application to Architecture*, J. B. Lippincott & Co., 1921. 70, 99, 110, 197

[18] P. Coleman and K. Singh, *RYAN: Rendering Your Animation Nonlinearly projected*, in Proceedings of the 3rd international symposium on Non-Photorealistic Animation and Rendering, 2004, pp. 129–156. 25

[19] J. P. Collomosse and P. M. Hall, *Cubist Style Rendering from Photographs*, IEEE Transactions on Visualization and Computer Graphics, 9:4 (2003), pp. 443–453. 25

[20] W. Cowan, *Rendering with Limited Means*. `http://www.cgl.uwaterloo.ca/~wmcowan/research/art/emse.html` (accessed December 2012), 1995. 4

[21] H. S. M. Coxeter, *Projective Geometry*, University of Toronto Press, 1974. 226

[22] A. Criminisi, M. Kemp, and A. Zisserman, *Bringing Pictorial Space to Life: Computer Techniques for the Analysis of Paintings*, in Proceedings of Computers and the History of Art (CHArt), 2002, pp. 1–33. 26, 168

[23] T. Davis, *Homogeneous Coordinates and Computer Graphics*. `http://www.geometer.org/mathcircles/homogen.pdf` (accessed December 2012). 116, 200

[24] C. DeCoro, F. Cole, A. Finkelstein, and S. Rusinkiewicz, *Stylized Shadows*, in Proceedings of the 5th international symposium on Non-Photorealistic Animation and Rendering, 2007, pp. 77–83. 23

[25] P. della Francesca, *De Prospectiva Pingendi (On Perspective for Painting)*, Casa Editrice Le Lettere, 1994. 46, 55

[26] J. S. Donath, *Structure Video and the Construction of Space*, in Proceedings of IS&T/SPIE's Symposium on Electronic Imaging, 1997. 24

[27] J. Dorsey and L. McMillan, *Computer Graphics and Architecture: State of the Art and Outlook for the Future*, ACM SIGGRAPH Computer Graphics, Newsletter, 32:1 (1998), pp. 45–48. 3

[28] W. V. Dunning, *Changing Images of Pictorial Space: A History of Spatial Illusion in Painting*, Syracuse University Press, 1991. 51

[29] F. Durand, *The Art and Science of Depiction*. `http://people.csail.mit.edu/fredo/gbRealisme.pdf` (accessed December 2012), 2000. 44

[30] ——, *An Invitation to Discuss Computer Depiction*, in Proceedings of the 2nd international symposium on Non-Photorealistic Animation and Rendering, 2002, pp. 111–124. 4

[31] S. Y. Edgerton, Jr., *Alberti's Perspective: A New Discovery and a New Evaluation*, The Art Bulletin, 48:3/4 (1966), pp. 367–378. 54

[32] ——, *The Renaissance Rediscovery of Linear Perspective*, Basic Books, 1975. 1, 56

[33] K. Elam, *Geometry of Design: Studies in Proportion and Composition*, Princeton Architecture Press, 2001. 35

[34] J. Elkins, *Perspective in Renaissance Art and in Modern Scholarship*, PhD thesis, University of Chicago, 1989. 38

[35] ——, *Art History and the Criticism of Computer-Generated Images*, Leonardo, 27:4 (1994), pp. 335–42. 3

[36] ——, *The Poetics of Perspective*, Cornell University Press, 1994. 4, 42, 43, 44, 45, 49, 53

[37] J. V. Field, *The Invention of Infinity: Mathematics and Art in the Renaissance*, Oxford University Press, 1997. 1, 45, 46, 47, 55, 59, 164, 169, 170, 194

[38] E. Fourquet, *Composition in Perspectives*, in Proceedings of the 4th Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging, 2008, pp. 9–16. 27, 73

[39] ——, *Learning About Shadows from Artists*, in Proceedings of the 6th international conference on Computational Aesthetics in Graphics, Visualization and Imaging, 2010, pp. 107–114. 73

[40] E. Fourquet, W. Cowan, and S. Mann, *On the Empirical Limits of Billboard Rotation*, in Proceedings of the 4th symposium on Applied Perception in Graphics and Visualization, 2007, pp. 49–56. 21, 173, 193

[41] ——, *Geometric Displacement on Plane and Sphere*, in Proceedings of Graphics Interface, 2008, pp. 193–202. 193

[42] I. S. Franke, S. Pannasch, J. R. Helmert, R. Rieger, R. Groh, and B. M. Velichkovsky, *Towards Attention-Centered Interfaces: An Aesthetic Evaluation of Perspective with Eye Tracking*, ACM Transaction on Multimedia Computing, Communications, and Applications, 4:3 (2008), pp. 18:1–18:13. 2, 3, 24

[43] W. Fulton, *Young Tableaux: With Applications to Representation Theory and Geometry*, no. 35 in London Mathematical Society Student Texts, Cambridge University Press, 1997. 157

[44] P. Gawthrop, *Perspective and Pictorial Space*. `http://www.lightspacewater.net/Tutorials/OCA/Perspective_Space.pdf` (accessed December 2012), February 2010. 32

[45] L. Geatti and L. Fortunati, *The Flagellation of Christ by Piero della Francesca: A Study fo Its Perspective*, in Visual Mind: Art and Mathematics, M. Emmer, ed., MIT Press, 1992, pp. 207–214. 169, 170

[46] M. Glueck, K. Crane, S. Anderson, A. Rutnik, and A. Khan, *Multiscale 3D Reference Visualization*, in Proceedings of the 2009 symposium on Interactive 3D graphics and games, 2009, pp. 225–232. 18

[47] E. H. Gombrich, *Art and Illusion: A Study in the Psychology of Pictorial Representation*, Princeton University Press, 1960. 2, 4

[48] B. Gooch and A. Gooch, *Non-Photorealistic Rendering*, A K Peters, Ltd, 2001. 3, 24

[49] B. Gooch, E. Reinhard, C. Moulding, and P. Shirley, *Artistic Composition for Image Creation*, in Proceedings of the 12th Eurographics Workshop on Rendering Techniques, 2001, pp. 83–88. 32

[50] A. Grynberg and G. Ward, *Conference simulation*. `http://www.anyhere.com/gward/pixformat/tiffluvrend.html` (accessed on August 2011). 31

[51] H. Hans, *Search for the Real*, MIT Press, revised ed., 1967. 31

[52] P. Heckbert, *Fundamentals of Texture Mapping and Image Warping*, Master's thesis UCB/CSD 89/516, Computer Science Division, University of California, Berkeley, June 1989. 247, 251

[53] P. Hendy, *Piero della Francesca and the Early Renaissance*, Macmillan, 1968. 38

[54] A. Hertzmann, *Painterly Rendering with Curved Brush Strokes of Multiple Sizes*, in Proceedings of SIGGRAPH 98, Annual Conference Series, 1998, pp. 453–460. 24

[55] D. Hockney, *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters*, Thames & Hudson Ltd, 2001. 50, 52

[56] D. Hoiem, A. A. Efros, and M. Hebert, *Automatic Photo Pop-up*, ACM Transactions on Graphics, 24:3 (2005), pp. 577–584. 20

[57] D. Hoiem, A. A. Efros, and M. Hebert, *Putting Objects in Perspective*, International Journal of Computer Vision, 80:1 (2008), pp. 3–15. 21

[58] Y. Horry, K.-I. Anjyo, and K. Arai, *Tour Into the Picture: Using a Spidery Mesh Interface to Make Animation from a Single Image*, in Proceedings of SIGGRAPH 97, Annual Conference Series, 1997, pp. 225–232. 20

[59] J. Huang, *Image-Based Relighting*, Master's thesis, University of Waterloo, 2010. 23

[60] K. Hulsey, *Adobe Illustrator & Photoshop Tutorials & Lessons*. `http://www.khulsey.com/student.html` (accessed on December 2012). 11

[61] T. Igarashi, S. Matsuoka, and H. Tanaka, *Teddy: A Sketching Interface for 3D Freeform Design*, in Proceedings of SIGGRAPH 99, Annual Conference Series, 1999, pp. 409–416. 19

[62] J. Jia, J. Sun, C.-K. Tang, and H.-Y. Shum, *Drag-and-Drop Pasting*, ACM Transaction on Graphics, 25:3 (2006), pp. 631–637. 22

[63] T. E. D. John R. Pani, Julia H. Chariker and N. Johnson, *Acquiring new spatial intuitions: Learning to reason about rotations*, Cognitive Psychology, 51:4 (2005), pp. 285–333. 18

[64] T. D. C. Kaufmann, *The Perspective of Shadows: The History of the Theory of Shadow Projection*, Journal of the Warburg and Courtauld Institutes, 38 (1975), pp. 258–287. 72

[65] M. Kemp, *The Science of Art: Optical Themes in Western Art from Brunelleschi to Seurat*, Yale University Press, 1992. 1, 33, 34, 35, 36, 40, 42, 45, 46, 48, 55, 97, 162, 164, 169, 170, 247, 257

[66] W. B. Kerr and F. Pellacini, *Toward Evaluating Lighting Design Interface Paradigms for Novice Users*, ACM Transaction on Graphics, 28:3 (2009), pp. 26:1–26:9. 22

[67] D. E. Knuth, *The Art of Computer Programming*, vol. 3, Sorting and Searching, Addison-Wesley, 1973. 157

[68] M. Kubovy, *The Psychology of Perspective and Renaissance Art*, Cambridge University Press, 1986. 48

[69] M. A. Lavin, *Piero della Francesca's Baptism of Christ, with an Appendix by B. A. R. Carter*, Yale University Press, 1981. 37, 38, 166

[70] M. H. R. Leyssen, S. Linsen, J. Sammartino, and S. E. Palmer, *Aesthetic preference for spatial composition in multiobject pictures*, i-Perception, 3:1 (2012), pp. 25–49. 32

[71] L. Liu, R. Chen, L. Wolf, and D. Cohen-Or, *Optimizing photo composition*, Computer Graphics Forum, 29 (2010), pp. 469–478. 32

[72] E. Loran, *Cézanne's Composition: Analysis of His Form with Diagrams and Photographs of His Motifs*, University of California Press, 2006. 25

[73] K. Mashio, K. Yoshida, S. Takahashi, and M. Okada, *Automatic Blending of Multiple Perspective Views for Aesthetic Composition*, in Proceedings of the 10th international conference on Smart graphics, 2010, pp. 220–231. 25

[74] R. J. Masters, *Computer Synthesis of Anamorphic Projection Systems*, Computer Graphics and Art, (1978), pp. 10–19. 40

[75] B. Meier, *Computers for Artists Who Work Alone*, ACM SIGGRAPH Computer Graphics, Newsletter, 33:1 (1999), pp. 50–51. 14, 18

[76] B. J. Meier, *Painterly Rendering for Animation*, in Proceedings of SIGGRAPH 96, Annual Conference Series, 1996, pp. 477–484. 24

[77] J. Meyer, *Multiperspective Collages*. ACM SIGGRAPH 2005 Sketches, Article No. 2, 2005. 25

[78] T. Mical, *The Origins of Architecture, After De Chirico*, Art History, 26:1 (2003), pp. 78–99. 184

[79] M. Mills, *Image synthesis: Optical identity or pictorial communication*, in Proceedings of Graphics Interface, 1985, pp. 303–308. 3

[80] M. E. Newell, R. G. Newell, and T. L. Sancha, *A New Approach to the Shaded Picture Problem*, Proceedings of the ACM National Conference, (1972), pp. 443–450. 130

[81] J. North, *The Ambassadors' Secret*, Hambleton & London, 2002. 38, 40, 53

[82] P. Olmer, *Perspectives Artistiques*, Plon, Paris, 1949. 48

[83] S. E. Palmer and S. Guidi, *Mapping the perceptual structure of rectangles through goodness-of-fit ratings*, Perception, 40:12 (2011), pp. 1428–1446. 81

[84] F. Pellacini, P. Tole, and D. P. Greenberg, *A User Interface for Interactive Cinematic Shadow Design*, ACM Transaction on Graphics, (2002), pp. 563–566. 23

[85] M. A. Penna and R. R. Patterson, *Projective Geometry and Its Applications to Computer Graphics*, Prentice-Hall, 1986. 200

[86] M. A. Peterson, B. Gillam, and H. Sedgwick, eds., *In the Mind's Eye: Julian Hochberg on the Perception of Pictures, Films, and the World*, Oxford University Press, 2007. 4, 19

[87] L. A. Piegl, *Ten challenges in computer-aided design*, Computer-Aided Design, 37:4 (2005), pp. 461–470. 18

[88] P. Poulin and A. Fournier, *Lights from Highlights and Shadows*, in Proceedings of the 1992 symposium on Interactive 3D graphics, 1992, pp. 31–38. 22

[89] T. Puttfarken, *The Discovery of Pictorial Composition: Theory of Visual Order in Painting 1400–1800*, Yale University Press, 2000. 34, 46, 51, 53

[90] J. B. Rapp, *A geometrical analysis of multiple viewpoint perspective in the work of Giovanni Battista Piranesi: an application of geometric restitution of perspective*, The Journal of Architecture, 13:6 (2008), pp. 701–736. 26

[91] M. Reid, *Undergraduate Algebraic Geometry*, no. 12 in London Mathematical Society Student Texts, Cambridge University Press, 1988. 113, 116

[92] T. Ritschel, K. Templin, K. Myszkowski, and H.-P. Seidel, *Virtual Passepartouts*, in Proceedings of the symposium on Non-Photorealistic Animation and Rendering, 2012, pp. 57–63. 20

[93] A. Rivers, F. Durand, and T. Igarashi, *3D Modeling with Silhouettes*, ACM Transaction on Graphics, 29:4 (2010), pp. 109:1–109:8. 19

[94] A. Rivers, T. Igarashi, and F. Durand, *2.5D Cartoon Models*, ACM Transaction on Graphics, 29:4 (2010), pp. 59:1–59:7. 19

[95] S. K. Schmidt R., Khan A. and K. G., *Analytic Drawing of 3D Scaffolds*, ACM Transaction on Graphics, 28:5 (2009), pp. 149:1–149:10. 18, 27

[96] T. K. Sharpless, B. Postle, and D. M. German, *Pannini: A New Projection for Rendering Wide Angle Perspective Images*, in Proceedings of the 6th international conference on Computational Aesthetics in Graphics, Visualization and Imaging, 2010, pp. 9–16. 26

[97] A. Shesh, A. Criminisi, C. Rother, and G. Smyth, *3D-aware Image Editing for Out of Bounds Photography*, in Proceedings of Graphics Interface, 2009, pp. 47–54. 20, 21

[98] K. Singh, *A Fresh Perspective*, in Proceedings of Graphics Interface, 2002, pp. 17–24. 25

[99] J. Stillwell, *The Four Pillars of Geometry*, Undergraduate Texts in Mathematics, Springer, 2005. 6, 7

[100] D. G. Stork and Y. Furuichi, *Computer graphics synthesis for inferring artist studio practice: An application to Diego Velázquez's Las Meninas*, in Proceedings of SPIE 7238, The Engineering Reality of Virtual Reality, 2009, pp. 723806:1–723806:10. 26

[101] I. Sutherland, *Sketchpad: A Man-Machine Graphical Communication System*, in Proceedings of the 1963 Spring Joint Computer Conference, vol. 23 of AFIPS Conference Proceedings, 1963, pp. 329–346. 18

[102] R. Talbot, *Speculations on the Origins of Linear Perspective*, Nexus Network Journal, 5:1 (2003), pp. 64–98. 75, 162, 163, 170, 176, 244

[103] ——, *Ambiguity and the development of Linear Perspective*, Tracey: Ambiguity, (2006), pp. 1–13. 67

[104] B. Thomas, *Geometry in Pictorial Composition*, Oriel Press, 1969. 36

[105] Toon Boom Animation Inc. `http://www.toonboom.com` (accessed December 2012). 21

[106] G. Vasari, *Vies des artistes*, Essai (broché), Grasset Collection Cahiers Rouges, 2007. 44, 55

[107] J. Willats and F. Durand, *Defining Pictorial Style: Lessons from Linguistics and Computer Graphics*, Axiomathes, 15:3 (2005). 6

[108] J. Yu and L. McMillan, *A Framework for Multiperspective Rendering*, in Proceedings of the 15th Eurographics workshop on Rendering Techniques, 2004, pp. 61–68. 25

[109] M. Zavesky, J. Wojdziak, K. Kusch, D. Wuttig, I. S. Franke, and R. Groh, *An Individual Perspective–Perceptually Realistic Depiction of Human Figures*, in Proceedings of the international joint conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2011, pp. 313–319. 28

[110] R. C. Zeleznik, K. P. Herndon, and J. F. Hughes, *SKETCH: An Interface for Sketching 3D Scenes*, in Proceedings of SIGGRAPH 96, Annual Conference Series, 1996, pp. 163–170. 19

[111] L. Zelnik-Manor and P. Perona, *Automating Joiners*, in Proceedings of the 5th international symposium on Non-Photorealistic Animation and Rendering, 2007, pp. 121–131. 25

[112] Y. Zheng, X. Chen, M.-M. Cheng, K. Zhou, S.-M. Hu, and N. J. Mitra, *Interactive Images: Cuboid Proxies for Smart Image Manipulation*, ACM Transactions on Graphics, 31:4 (2012), pp. 99:1–99:11. 20

[113] D. Zorin and A. H. Barr, *Correction of Geometric Perceptual Distortions in Pictures*, in Proceedings of SIGGRAPH 95, Annual Conference Series, 1995, pp. 257–264. 24, 33, 44