Computational Aesthetics in Graphics, Visualization, and Imaging (2008)
P. Brown, D. W. Cunningham, V. Interrante, and J. McCormack (Editors)

Composition in Perspectives

Elodie Fourquet[†]

David R. Cheriton School of Computer Science
University of Waterloo, Canada

Abstract

Composition is a key element of image aesthetics. However composition is hard to control when working in three dimensions to create a two dimensional image. A framework that derives perspective from a planar pattern is proposed and implemented. The third dimension is elevated from a tiled floor into a planar square pattern. Key points on the image allow users to modify the spatial geometry of the scene. Thus, this paper presents a new view on perspective, where there is no concrete third dimension, but where the third dimension is inferred from lines and points in the image plane from which apparent depth relationships of the scene are constructed.
In describing the framework, the computational relation between elements such as vanishing point, distance points and floor lines inside the geometric grid, are exposed to demonstrate the characteristics of building a realistic, yet, composed, image based on the practices of Renaissance painters.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation J.5 [Computer Applications]: Arts and Humanities K.3.m [Computer and Education]: Miscellaneous

1. Introduction

Seeing realistic art as a two dimensional perspective projection of a three dimensional scene is a naïve view of artistic practice. In effect, the best realistic art is the result of two dimensional patterns in the image plane, an 'accidental' result of three dimensional placement combined with perspective projection. Thus, image plane composition is an essential element of the expressiveness of an image. The emotional difference between family snapshots and artistic photography lies in the composition of the latter. Before taking the photograph the artist carefully arranges the scene and the camera, positioning each element so that projection produces a significant set of two dimensional relationships.

Painters usually work on a flat surface. Medieval artists emphasized composition on the painted surface above all. They arranged the elements of a picture to form geometric patterns with symbolic meaning. For example, an equilateral triangle often organized pictures showing the Trinity. Later, Renaissance artists combined this practice with geometric perspective, which gives a good sense of 3D space. Largely developed through trial and error, apparent 3D in the Renaissance depends on the relationships of sets of lines, such as the edges of a cube or of tiles on a floor, placed using 3D construction lines. How to reduce dimension was learned by generalizing picture plane constructions, a practice that formed the mathematics of projective geometry. How artists discovered perspective is important, to be sure, but more important is how they learned to master the simultaneous compositional demands of 2D and 3D.

Unfortunately, despite the importance of 2D composition, contemporary 3D modelling software provides little support for it. Within modelling software, the user works from a 3D scene definition to its 2D projection. That is, the user first edits 3D objects and arranges them in a virtual space and, after positioning the camera, receives a projected image. Thus, iteration between 3D edits and re-examination of the image is the work-style, providing only indirect control of 2D composition. With enough practice many users master the skills needed to visualize the effect on 2D composition of object manipulation in 3D, but most artists reject 3D modelling. They prefer to use tools like Adobe Photoshop or Illustrator, where they have direct control of 2D composition, but at the price of rudimentary support for composition in depth.

To understand the problem consider how artists work.

[†] efourque@cgl.uwaterloo.ca

Representational art usually begins with construction lines that organize the picture surface. If the portrayal of space is important, construction lines for perspective projection are included, harmonized with the remainder of the composition. In extreme cases, such as the *School of Athens* or a cruise ship [Hus], creating the construction lines is the most time-consuming aspect. My hypothesis is that artists work like this because it provides them with the control they require over the final image, which is simply not available from 3D tools. Thus, I have created a prototype tool for supporting image creation, which is a testbed for investigating how to integrate control over 2D and 3D composition.

Art of the Italian Renaissance, the 'golden age of perspective', in its employment of optical realism, is close to the spirit of computer graphics. Yet these mathematician/artists used perspective with a freedom that is currently impossible in computer graphics, finding an intermediate ground between rigid unitary perspective and the 'anything goes' realm of exotic perspective. According to art theorist James Elkins, Renaissance perspective can be viewed as a collage of constructions [Elk94]. Especially in early Renaissance, when perspective methods were developed and simultaneously mixed with Medieval picture plane composition, artists took a pluralist approach to perspective. By exploring these methods, which are taught in art schools even today, I am trying to identify the geometric constructions an artist needs to combine 2D and 3D composition.

Recently, Richard Talbot devised a grid of squares that combines image plane and depth geometry, and showed that such a grid can represent the composition of some early Renaissance paintings [Tal03]. Subdivisions of squares in a planar grid simultaneously organize the picture plane and the pictorial space, which is anchored by a tiled floor. That is, objects that require a well-defined location in pictorial space are located on the floor, which provides a strong illusion of depth. Strong integration of the grid with the floor perspective makes it possible to develop a volume that is coherent with the picture plane composition.

This paper describes a composition tool I have implemented that uses Talbot's construction, and extends it to allow placement of volumes and images. The geometry is purely two dimensional, giving the illusion of depth with perspective constructions used in the early Renaissance. Most of the paper is devoted to explaining the geometrical relationships and why they are important. That I use Talbot's construction is not to imply that I consider it to be the final word in picture geometry: many geometries are possible. If my research pushes artists to devise better ones then my goal has been achieved.

2. Previous Work

Much research has been done to reconstruct a 3D scene from one or more 2D projections of it. The reconstruction accuracy differs according to the goal, from rough approximations that permit quick and easy navigation, to accurate estimations that develop precise measurements of the scene from projections of it. For example, Horry et al. showed how to reconstruct, with the assistance of the user, the volume and foreground present in a photograph [HAA97]. A spidery mesh is mapped to the perspective of the photograph, with planes attached to the ground and to facades of objects that enclose the pictorial volume, and with billboards representing foreground objects. While elements of this system are similar to elements of mine — his facades are my walls, his billboards are my panels — his intent is different, to recompose rather than to compose.

The research of Criminisi is a formal approach to 3D scene reconstruction, using projective geometry and the geometric characteristics of the scene to reconstruct a three-dimensional model of the scene, without requiring knowledge of camera parameters [Cri01]. Criminisi's examples include the reconstruction of scenes in Renaissance art. While this work analyses Renaissance perspectives mathematically and derives the structures they depict, including possible projective inconsistencies, it neglects to consider the effects of picture plane composition within realistic depiction.

Research on the difference between computer graphics rendering and artistic rendering is more relevant. Examining Raphael's *School of Athens*, Zorin and Barr noticed that an optically incorrect artist's rendering may be more visually effective than a optically correct computer rendering [ZB95]. In such instances multiple perspectives have been introduced to enhance computer rendering. Building on multiple perspectives researchers created rendering systems with non-linear projection as a break from the unitary perspective of classical computer graphics [AZM00, Sin02]. This research, which introduces flexible perspective as a feature of rendering, is a complement to mine, which innovates in image composition, that is to say modelling.

Recent work on picture plane composition used artistic heuristics to control scene geometry [GRMS01]. An algorithm computes a format, viewpoint and layout for the image of a single 3D object, based on artistic criteria. This work, which makes use of heuristics from photography to select camera position, does not handle complex scenes and relies on the traditional unitary perspective of computer graphics.

Thus, the research described in this paper addresses a problem that is new to computer graphics. It relies on recent research in art history and education [Tal03], which does not touch on computer graphics. Talbot's work is a contribution to the large field of pictorial geometry and the symbolism of art. For lack of space I do not attempt to survey this copious literature, particularly because Talbot's specific geometry was chosen because it is easily adapted to computer graphics, and not because it is unique.

3. Composed Perspective Framework

Working on a flat surface, artists perceive and interact with the projection of a scene as they create it. The process is fundamentally different from that of a user developing a scene with a modelling software, where each object is fully modelled in three dimensions and positioned in the scene before being perspective projected onto the picture plane. In 3D modelling the picture plane composition is controlled indirectly and the results known *post hoc*. In the practice of realistic artists, by contrast, each line, as it is drawn, belongs simultaneously to an object and to its projection, so that the composition develops incrementally during the construction of the image. Thus, in a two dimensional modelling process that supports an artistic work-style the user will create and transform construction lines while developing the picture. A method for doing so is found in artistic practice, and how artists control depth.

In early paintings that use perspective to represent 3D, artists often relied on a tiled floor to frame interior space. Such floors have a simple geometry, parallel lines along tile edges and through tile corners, coherent over large areas. In depth the tiles make simple patterns in the picture plane, which are easy to construct using only a straightedge. Indeed, the floor solves two hard problems of perspective construction: where to place an object on the picture plane is indicated by tile location, how much to scale the object is indicated by tile size. Thus, in early Renaissance painting we see human figures, which appear to float in Medieval painting, firmly anchored on tiled floors.

Using patterns of tiled floors combined with picture plane geometry, the artist can simultaneously control 2D and 3D composition. This observation is not new: many such frameworks exist, created to teach perspective drawing, and given authority by analyses of art history. That Renaissance artists discovered linear perspective by noticing construction lines in the picture plane that create an illusion of depth is attractive and plausible, though neither provable nor disprovable. If so, the role of Brunelleschi and Alberti is less to have discovered perspective than to have formalized it.

My prototype tool uses the geometry proposed by Talbot [Tal03]. Before describing that geometry for 2D composition, I define several important terms, then show how perspective projection of the tiled floor fits naturally within the 2D geometry, setting the proportions needed for depicting the rest of the scene. The organization of space completed, I then describe how to mass volumes on the picture place and in depth.

3.1. Definitions

Currently my prototype supports simple perspective constructions, which maximize balance and symmetry. The constructions are based on several geometric concepts, which are defined below.

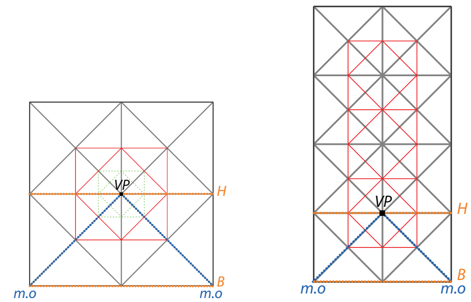
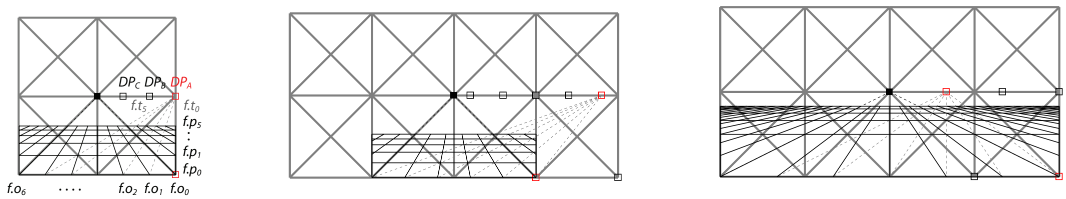


Figure 1: *Two configurations of grid of squares providing the image plane geometry in my framework. The grids show the recursed inscribed pattern. The labels and their dashed lines are only present as guide marks for the reader.*

- Horizon line, H : The horizontal line containing the vanishing point. Pictures are normally hung with this line at the height of the viewer's eyes.
- Baseline, B : The horizontal line defining the bottom of the picture.
- Floor orthogonals, $f.o$: Lines along tile edges perpendicular to the picture plane in 3D. Intersections of floor orthogonals with the baseline define horizontal tile divisions.
- Vanishing point, VP : The point on the horizon line where floor orthogonals meet.
- Distance point, DP : A point on the horizon line where diagonal lines in the tiled floor meet. The position of the distance point with respect to the vanishing point and the edge of the picture defines the distance from the eye to the picture. Of the two distance points only one is needed to define the construction.
- Main orthogonals, $m.o$: The two floor orthogonals from the vanishing point to the bottom corners of the picture.
- Floor parallels, $f.p$: Lines along tile edges parallel to the picture plane in 3D. Horizontal lines at the edges of tiles define the foreshortening at different depths.
- Distance point diagonal lines or floor transversals, $f.t$: Lines from a distance point to intersections of the baseline with the floor orthogonals. Intersections of floor transversals with the main orthogonals locate the floor parallels on the picture plane.

3.2. The Image Plane Grid

Talbot provides a construction grid of square subdivisions that organizes the image on the picture plane [Tal03]. Thus, in my compositional tool, the basic pattern used (on the left of Figure 1) is a square and its reflection lines. A diamond is inscribed in the square, then a square (shown in red) in the diamond and so on. All inscribed squares and diamonds share a set of reflection lines. On the left of Figure 1, the basic pattern is recursed twice. A more complex compositional



a: 1x1 grid, *DP* on frame | b: 2x1 grid, *DP* outside frame | c: 2x1 grid, *DP* close to *VP*

Figure 2: *Different floor configurations, representing various views.*

grid is shown on the right of Figure 1: a vertical juxtaposition of two basic patterns recursed once provides a grid suitable for the nave of a church. Sub-squares are added to bridge inscribed squares.

When composing a realistic image the artist first sets the eye level and view direction by determining the position of the vanishing point on the horizon line. In my current implementation the vanishing point is placed in the centre of the square in the middle of the lowest row of squares. Thus, the horizon line is the horizontal reflection line with the vanishing point at its centre. The basic square defines the foreshortening of the floor tiles. The first inscribed square defines the walls, and the second inscribed square provides locations at which objects can be placed in accord with the 2D geometry.

The grid pattern may exceed the image width, which places construction points outside the picture, a common practice. The image width is adjusted by moving the main orthogonal, *m.o*, along the baseline to corners of reflection lines in the basic pattern(s). In Figure 2, rectangles on the baseline are possible anchor points for the picture corner. In early Renaissance art the distance point was usually placed on the edge of the painting, but distance points beyond the edge are also common.

3.3. The Tiled Floor in Perspective

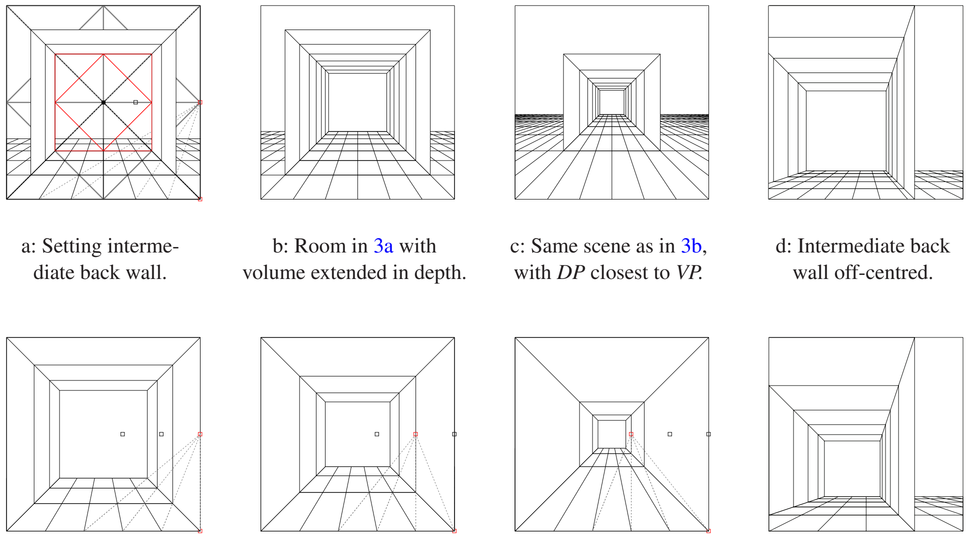
In 3D a tiled floor has two sets of parallel lines at right angles, in a horizontal plane. Usually the view is chosen so that the lines are perpendicular and parallel to the picture plane. Projected to the picture plane, the floor is a converging set of lines, *f.o*, meeting at the vanishing point, crossed by a set of lines, *f.p*, parallel to the horizon. All tiles have the same width and the floor orthogonals intersect the baseline at regular intervals. For example, in Figure 2, the floors have either 5 or 6 tiles on the baseline.

The spacing of the floor parallels is the quantitative recession in depth of the tiles, which defines the position of the viewpoint. Many perspective constructions exist for placing the floor parallels [Kem92]. All determine the floor geometry using intersections of construction lines. The floor orthogonals are fixed by the number of tiles. There are many possible spacing of the floor parallels, depending on the position of the viewpoint. As important as the verisimilitude of the projection is the harmonic pattern of the floor parallels, which is visually pleasing.

Art historians have long analyzed and debated the writings of Alberti [Alb35], questioning the extent to which theory influenced the construction lines in Renaissance paintings [Edg66, Kem92]. Talbot's picture plane grid combines well with current understanding of Alberti's floor construction. The vanishing point, *VP*, and distance point, *DP*, lie on a reflection line of the main square. The distance between them is the distance of the eye from the picture plane, and of the scene portrayed, as Alberti suggests.

The *DP* is the point of convergence of 45° lines that lie along the tile diagonals, which are called floor transversals, *f.t*. The floor transversals connect tile corners on the baseline to the *DP*. Thus, on the baseline, each floor orthogonal, *f.o*, connects to a floor transversal, *f.t*. In the simplest case, the *DP* is located at the edge of the painting as in Figure 2a. The angle of view represented in the painting is 90° and the viewer's eye is half the painting width from the painting when seeing it *'correctly'*. The floor parallels, *f.p*, pass through the intersections of the main orthogonals, *m.o*, with the successive floor transversals, *f.t*, defining the foreshortening of the tiled floor. From the nearest row of tiles, on the baseline, to the farthest row, the transversal segments decrease in length and intersect the *m.o* farther away from the picture corner. Specifically, the first *f.t* lies along the edge of the picture, intersecting the *m.o* on the baseline, which is the first floor parallel line, the null depth. The second *f.t* intersects the *m.o* at the point defining the first tile depth, and so on. That this construction coincides exactly with the predictions of linear perspective can be shown elegantly by a geometric proof based on the distance point method [Kem92], a proof with which artists are comfortable.

Moving the *DP* closer to the *VP* moves the viewing point closer to the picture plane. Conversely, moving the *DP* beyond the edge of the picture moves the eye farther from the picture plane. Indeed, as the eye point moves closer the nearest tiles get longer in the picture plane, while the farther ones get shorter. Figure 2 shows examples of the floor with three different distance points. Figure 2a is the usual case with



a: Setting intermediate back wall.

b: Room in 3a with volume extended in depth.

c: Same scene as in 3b, with *DP* closest to *VP*.

d: Intermediate back wall off-centred.

e: Cube looking volume.

f: *DP* in between of *DP's* of 3e and 3g.

g: Tunnel looking volume

h: Same as in 3d with wider tile on baseline.

Figure 3: *A variety of pictorial volumes.*

the *DP* at the edge of the painting. The six *f.t* intersect the right *m.o* at increasingly distant parallel lines. Figure 2b has a floor configuration close to Figure 2a, but the *DP* is outside the picture. The edges of the picture coincide with the vertical reflection lines of the two basic patterns. The longer distances of *DP* from the transversal intersections with the baseline reduce the foreshortening extending the apparent depth of the floor. Figure 2c shows the same floor as the one in Figure 2a, but on a 2×1 grid and with a *DP* relatively much closer to the *VP*. This configuration creates visual tension owing to the contrast between deep tiles at the baseline and a crowd of indistinguishable tiles farther back. The pictorial space has a fish-eye quality, but without the hyperbolic distortion resulting from a finite-sized lens. As in Talbot's construction the *DP* always lies directly above an intersection of a *f.o* with the baseline, which retains geometric harmony and relaxes tension. The construction succeeds when the *DP* is moved to other locations, but the geometry acquires undesirable tension.

The construction creates as many tiles receding in depth as there are tiles across the baseline. This is, however, an artificial restriction: the baseline can be extended beyond the grid to produce an unlimited number of tiles receding in depth. In exploring these capabilities within my implementation I quickly discovered what Renaissance artists learned by trial and error. First, having many more tiles in depth than across the picture plane makes pictorial space that looks unnaturally deep, and many fewer makes an unnaturally shallow space. Stages are frequently very shallow, designers handle such spaces with a hybrid perspective that increases the apparent depth. Second, placing the *DP* just outside the edge of the picture makes a pictorial space with more apparent depth.

Talbot's construction creates an intimate relationship between the grid pattern and lines defining the perspective geometry of the floor. The *f.o* defined by depth tile edges converge on the centre of a basic square; the diagonal *f.t*, running through tile corners, converge to *DP*s, located where the horizontal reflection line intersects the edge of the square; the *f.h* defined by horizontal tile edges are parallel to edges of the grid square and its horizontal reflection lines. Furthermore, the *f.h* recede from the baseline to the *VP* along the *m.o* in harmonic ratios. The *DP* moves discretely along the horizon, taking on positions directly above intersections of the *f.o* with the baseline, an implicit axis parallel to the edges of the grid.

3.4. The Objects on the Tiled Floor

3.4.1. Walls

The tiled floor provides a guide for locating objects in the scene. When the *DP* is on the picture edge, the base of the first inscribed square is in the tile half way back on the floor, where the width of the inscribed square is the distance between the *m.o*, as shown in Figure 3a. In effect, the first level inscribed squares are a plane between the *VP* and the baseline, half way back on the floor. (When the *DP* moves the

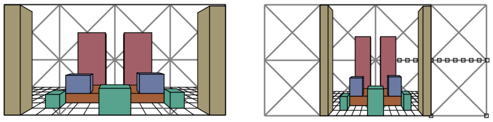


Figure 4: *Multiple blocks. The right image is produced from the left image, by changing the picture edge and with theDP outside the image.*

relationship between half way back on the floor and the inscribed square breaks, adding tension to the composition.)

Back and side walls may be constructed to close the pictorial space. The back wall starts as a temporary wall, half way back on the floor. It is a column of first level inscribed squares rising as high above the floor as the user desires. Side walls are then created, based on rays from the *VP* and passing through the top corners of the temporary back wall, the dashed lines in Figure 3a. Vertical lines are drawn up to the rays from points where each *f.p* in front of the temporary wall intersects the two ray segments of the intermediate wall that defines the side wall orthogonal lines on the floor. The wall vertical lines terminate on ray segments extended forward from the top corners of the intermediate wall, forming the structure of the side walls. To make a volume more convincing the side walls can be made opaque, hiding tiles behind the walls.

Figure 3a shows the volume in front of the temporary back plane. The temporary back wall is then moved to create the desired depth, the side walls being extended with it. Figure 3b shows the result. This configuration is constructed with the *DP* at the edge of the grid. Figure 3c shows the same scene with the *DP* as close to the *VP* as it can be placed. Figures 3e–3g show intermediate positions of the *DP*. The relatively cubic volume in Figure 3e is transformed to a tunnel in Figure 3g. The resulting back wall is smaller because the floor is raised and appears more steep. Figure 3g might be Figure 3e seen by a mouse. Finally, in all the images of Figure 3 a single centred first level inscribed square was used to set the intermediate back wall, except for Figures 3d and 3h where two side-by-side squares were used, the square to the left containing the *VP*. Such a choice of back wall makes the volume asymmetric. Figures 3d and 3h show the effect of changing the number of horizontal floor divisions. In Figure 3d, the tiles are more numerous and smaller in width, which makes a bigger looking back wall than in Figure 3h.

3.4.2. Blocks

The walls define the pictorial volume, seen from inside. Similarly, blocks, seen from outside, can be added to define volumes within the scene. Blocks allow the artist to plan the visual interaction of volumes occupied by objects in the scene, such as architectural content. Blocks rest on the floor, which fixes their depth firmly. They provide surfaces on which other elements can rest, and facilitate composing in height.

To build a block, the bottom face is defined as a rectangular array of floor tiles, after which the front face of the block is elevated to the desired height. Locations of corners defining the sides, back and top are calculated from the perspective construction. Specifically, the two intersections of the *m.o* with both the *f.p* at the front and at the back of the block are calculated. These points define two verticals, with their origin at each intersection. On the front vertical, the upper corners of the front of the block lie on lines passing through the *VP* that intersect the vertical back line at the foreshortened height of the back face. Using the main orthogonal as the reference for height diminution provides uniformity in height changes across the width of the picture.

Figure 4 shows several blocks standing on the floor and above other blocks. The image on the right shows how moving the *DP* and the picture edge changes the geometry and the volume of the blocks.

3.4.3. Panels

Finally, 2D image content can be mapped onto panels, which may be inserted within the other constructions. Thus, the picture can include the collages of constructions identified by Elkins in Renaissance paintings [Elk94]. Studio drawings as cartoons, for example, were transcribed onto walls for fresco using overlaid grids to scale the figures, which were positioned using construction lines that defined the composition. I use transparent panels with opaque figures to maintain proper occlusion. To make positioning with respect to construction lines easier, manipulated panels are made transparent.

Panels are inserted, moved and scaled to fit the compositional grid, the perspective elements, the floor or the blocks. The visible grid and the floor tiles allow exact placement of the panels by snapping onto lines or intersections. When gravity is present the bottoms of panels are fixed to the floor. In its absence panels float free of the floor and are anchored to points on the grid.

Figure 5 contains examples of images that were composed with my system. The one on the left reproduces Masaccio's, *Holy Trinity*, using constructions synthesized by Talbot. It illustrates how my framework makes effective use of his perspective construction. The others are also inspired by Renaissance paintings. The middle ones are based on *Madonna and child*, by Jan van Eyck. The lower one is close to the composition and perspective of the actual painting, with the Madonna centred in depth and resting in front of the base of the first level inscribed square. In the top image, which differs only in the *DP* being closer to the *VP*, the Madonna remains centred in depth, but she is well above the base of the inscribed square in the plane composition and seems remote from the viewer. The last column of pictures in Figure 5 is

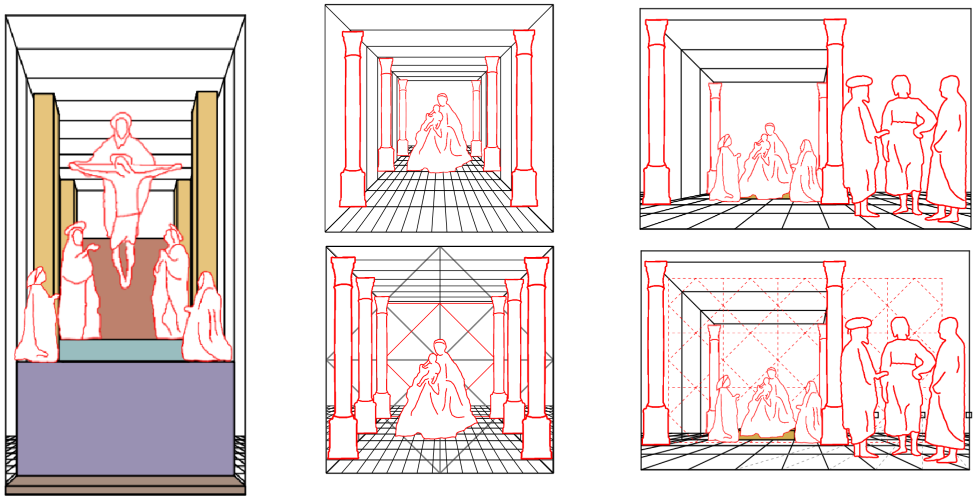


Figure 5: *Composition based on 'The Holy Trinity', c. 1428, of Masaccio. Composition based on 'Madonna and child', c. 1437, of Jan van Eyck. Hybrid composition inspired from perspective construction and composition in the 'Flagellation of Christ', after 1444, of Piero della Francesca and with other figures in the background.*

a composition based on the *Flagellation of Christ* by Piero della Francesca, but with different figures in the background: the flagellated Christ has been replaced by the Madonna and the child, while the torturers have been replaced by the pious women of Masaccio. The *Flagellation of Christ* is noteworthy because its off-centred perspective is unusual for its time. In my adaptation the *VP* is centred, but an uncentred room and a *DP* brought from the edge of the painting towards the *VP*, reproduces the unbalance. The bottom image shows the transversal lines from the *DP*, and the first level of inscribed squares, from which the intermediate wall was selected. It shows the scene relative to the planar geometry of its composition.

4. Implementation

The pictorial grid and the perspective construction described in Sections 3.2 and 3.3 were implemented in Java using Java2D. No 3D computation was performed. In this form I could easily explore many variations of pictorial space.

The objects described in Section 3.4 were implemented to explore the next stages of picture construction: limiting the pictorial space with walls, massing volumes within pictorial space, and placing pre-drawn images within pictorial space. Figures 1– 5 were all created with my prototype tool.

Since incorrect occlusion destroys the illusion of three dimensions, care is needed in the drawing order within and between elements, in particular for the blocks. Within each block, the front face must be drawn after the side faces, and the top face last, but only when the top is below the horizon. Otherwise the top is not drawn at all. Blocks are drawn from back to front, according to their positions on the floor. Blocks at the same distance are drawn from the edges of the picture toward the centre. When blocks overlap in depth, this simple algorithm needs to be extended to consider face categories. Also, as blocks can stand on top of other blocks, the bottom face of the upper block is raised to the height of the upper face of the lower block. Stacking is defined recursively, so blocks can be stacked to arbitrary levels. Stacking, however, further complicates the screen repainting algorithm, requiring several passes determined by the distance of the blocks bottom face to the horizon line.

5. Discussion and Future Work

The framework described in this paper was inspired by the artistic practice of Renaissance painters, with special attention paid to how they integrate the symbolic geometry of 2D composition with the organization of pictorial space. Integrating artistic practice into computer graphics tools is the best — if not the only — method for remedying the expressive shortfalls that are so often present in computer graphics images [Mil85, DM98].

The main contribution of this paper is my prototype implementation, which demonstrates that integrated 2D and 3D composition is possible. It does not claim that either the grid pattern or the perspective construction is optimal, only that the appearance of a 3D scene can be constructed from two dimensional patterns that provide interesting image composition. Indeed, my implementation shows that a tiled floor in perspective is an effective 'artistic compositional device' as Edgerton describes Alberti's floor construction [Edg66].

The framework can be developed further, especially in composition refactoring and content manipulation. It functions solely in terms of proportion. Thus, when something changes, such as the number of tiles in the width of the floor, elements anchored to the floor change in proportions and relative position. At present, a naïve update method is used, but examining the relationship between the image grid and the perspective can lead to improved conservation of balance and symmetry. Similarly, modifying view parameters such as the distance point or the picture edges should displace the panels so that they remain coherent within the new composition. Such extensions require a deeper study of proportion constraints in Renaissance art.

Another contribution is demonstrating how artists build perspective in the presence of image plane pattern and how 3D-like manipulation can occur within the compositional framework. Improvements that connect relative height and position in depth may further ease manipulation. I have shown how a one point perspective block can be raised from the floor. Adding block rotation will extend the framework, with convergence points of parallel lines in the side faces moving along the horizon line. Similarly, the framework currently considers only panels oriented parallel to the image plane. Panels with flat content may easily be mapped to differently oriented surfaces. More challenging are rotated panels that depict content having depth.

In the current implementation the final image contains multiple perspectives, the dominant perspective being determined by the floor and volume constructions. Local perspectives can be present in the panel images. However, nothing except the user's perception guarantees coherence between projections on the panels and the dominant perspective. Future work will include using constraints between recursive perspective grids to help users explore consistencies between multiple perspectives in an image.

Beyond my contributions to artistic computer graphics the integration of 2D composition with perspective has potential impacts on other areas of research. Three dimensional transformations occur increasingly in user interfaces and in visualization to produce content that is sufficiently information dense for displays of larger and smaller form factor. Here 3D is used, not to produce realism, but because it is an easily understood mental representation of information that is partly hidden, and thus can be manipulated easily. Because the display is planar, 2D composition is important, but must be combined with depth for intuitive access to partially displayed content. My framework has potential to enhance both aesthetic appeal and user efficiency. Easy perceptual organization, sought by artists for the sake of beauty, is also important for recognition and retrieval.

Acknowledgements – This work would not have been possible without the support and dedication of Bill Cowan and Stephen Mann, who are mentoring me as I learn to do research. Bill Cowan also helped me to make the prose of this paper more graceful.

In addition, I would like to thank all the members of the Computer Graphics Lab at the University of Waterloo for their comments and encouragement, and more specifically: Eoghan Sherry, Jeff Dicker, Martin Talbot.
This research was funded in part by NSERC and OGS.

References

[Alb35] ALBERTI L. B.: *De Pictura (first Latin edition, translated to English as On Painting)*. 1435.

[AZM00] AGRAWALA M., ZORIN D., MUNZNER T.: Artistic multiprojection rendering. In *Proceedings of Eurographics Workshop on Rendering Techniques* (2000), pp. 125–136.

[Cri01] CRIMINISI A.: *Accurate Visual Metrology from Single and Multiple Uncalibrated Images*. Springer-Verlag London Ltd, Distinguished Dissertation Series, 2001.

[DM98] DORSEY J., MCMILLAN L.: Computer graphics and architecture: State of the art and outlook for the future. *SIGGRAPH Computer Graphics 32* (1998), 45–48.

[Edg66] EDGERTON S. Y.: Alberti's perspective: A new discovery and a new evaluation. *The Art Bulletin 48*, 3/4 (1966), 367–378.

[Elk94] ELKINS J.: *The Poetics of Perspective*. Cornell University Press, Ithaca, 1994.

[GRMS01] GOOCH B., REINHARD E., MOULDING C., SHIRLEY P.: Artistic composition for image creation. In *Proceedings of Eurographics Workshop on Rendering Techniques* (2001), pp. 83–88.

[HAA97] HORRY Y., ANJYO K.-I., ARAI K.: Tour into the picture: using a spidery mesh interface to make animation from a single image. In *Proceedings of SIGGRAPH* (1997), pp. 225–232.

[Hus] HUSLEY K.:. http://khulsey.com/demo_1howto.html, accessed on 17 March, 2008.

[Kem92] KEMP M.: *The Science of Art: Optical Themes in Western Art from Brunelleschi to Seurat*. Yale University Press, New Haven, 1992.

[Mil85] MILLS M.: Image synthesis: Optical identity or pictorial communication. In *Proceedings of Graphics Interface* (1985), pp. 303–312.

[Sin02] SINGH K.: A fresh perspective. In *Proceedings of Graphics Interface* (2002), pp. 17–24.

[Tal03] TALBOT R.: Speculations on the origins of linear perspective. *Nexus Network Journal 5*, 1 (2003), 64–98.

[ZB95] ZORIN D., BARR A. H.: Correction of geometric perceptual distortions in pictures. In *Proceedings of SIGGRAPH* (1995), pp. 257–264.