NEW PERSPECTIVES FOR IMAGE SYNTHESIS

M. Beigbeder and V. Bourgoin
Ecole des Mines de Saint-Etienne
158, cours Fauriel
F42023 Saint-Etienne Cedex 2
tel: 77 42 01 23
email: {mbeig|bourgoin}@emse.fr

ABSTRACT

Various different perspectives are used in the representation world. In this paper, two kinds of perspective, the "fishbone" perspective, already existing before the Quattrocento, and another perspective coming from the concept of Cubism, are studied. We have tried to introduce these views of the mind in image synthesis. Two implementations are studied: a particle method and ray tracing.

Key Words: Computer graphics, Image generation, Viewing algorithms, Perspective transformation.

1 QUESTIONING CLASSICAL IMAGE SYNTHESIS

There are many different perspectives, choosing one of them reflects a way of interpreting the world. In image synthesis, conic perspective is the most widely used, and the rest being orthogonal ones. The camera obscura technique reveals conic perspectives (see Figure 1). Conic perspectives have been used in occidental art since the italian Renaissance. It was then, during the XV$^{\text{th}}$ century, that Brunelleschi found out a procedure which produced an illusion of space on a surface. Twenty years later in 1435, the architect Alberti [Alb35] mathematically formulated the empirical idea of Brunelleschi. The same idea gave birth to what we now call the ray tracing method.

It could be supposed that the camera obscura has been known since the Antiquity because frescoes using the conic perspective were found in some roman villas built in the first century B.C.. Let us trace the history of

Figure 1: The camera obscura.

the conic perspective initiated by the camera obscura:

- Brunelleschi transposed the principle of the camera obscura with a mirror[1];
- at the end of the XVI$^{\text{th}}$ century Giambattista della Porta improved the camera obscura by replacing the hole with a lens—the use of lenses had been developed during the XVII$^{\text{th}}$ century;
- the camera obscura became portable at the XVIII$^{\text{th}}$ century;
- in 1822, Nicephore Niepce used a sensitized plate, thus inventing photography[Neu67];
- the XX$^{\text{th}}$ century perpetuates this concept of realism in image synthesis.

Before the Renaissance pictures narrated and instructed the people. Drawing used to be akin to writing, it was a means to express a thought or an idea by re-presentation. The representation world might have been governed by cultural and religious rules. For example, egyptian art pictures any personage both front-view and side-face, which allows to see him from all sides. From the time of Greek art to that of mediaeval art, religion was the main driver of thought, it dictated the representation rules. For example, the personage sizes depend on their social importance.

After the Renaissance, the brunelleschian projection system was known as the only "realistic" model. It developped for several centuries until photography appeared, and perfectly achieved what painting was aiming at. Hence, photography, coming from the scientific domain, found its place within the representation world, a traditionally privileged field of art. This incited some painters to look for answers in science; and specially about light splitting. Chevreul's works on optical mixing are interpreted as an explosion of the shape unity into particles. But Impressionism preserves the single point of view of the Renaissance perspective. Cézanne began

Figure 2: Picasso.

to decompose volumes. Braque, an admirer of Cézanne, unfolded volumes, and initiated Cubism[Rub90]. The Cubism principle is to picture an object on a single plane from different points of view (see Figure 2).

The realism of the single point of view may be seen as a cultural convention. In fact, looking at an object means *scanning* it with *both* eyes, so there are two moving points of view. It is only the occidental mind abstraction that transforms the look into a single point of view[Pan87].

The realism of Cubism is just as legitimate as the realism of the camera obscura. The Renaissance perspective appears as a vision of the mind; so do the previous perspectives.

Our work consists in introducing these different visions of the mind into image synthesis. We have to interpret them both conceptually and mathematically, in order to question the so called *realistic* images. In each step of our work, we focused our attention on one aspect of the problem, and not on all the elements. Two scenes are used, the first one consists in a tiled floor and some ceil beams, the second one represents a beaver: they are shown in figure 3 rendered with the classical conic perspective.

2 INTERPRETING PAINTINGS

2.1 Observations

In a lot of paintings, it is possible to find representations of graphical references such as ceilings or tiled floors. These objects are reg-

[1] It should be noticed that some art historians don't agree with this point of view, Damisch for example

Figure 3: Two scenes rendered with the classical perspective.

Figure 5: Mural painting Ist century.

ular in that their components share a single size in some directions and are regularly disposed. Different perspectives can be revealed through this regularity.

A ceiling with parallel equidistant beams, projected with a brunelleschian perspective, with the projection direction parallel to them, draws lines which converge on the center of the "screen" (see Figure 4).

One of the most wide-spread perspective before Renaissance is the "fishbone" perspective. Its name come from the drawing of the parallel beams projection. Each couple of symmetrical beams intersect on a point, and all of these points are vertically aligned (see Figure 5). Though the order of these points may vary from one image to another (see Figures 6 and 7).

In the following paper, we shall first study how to move the vanishing points, and then we shall try to recreate broken lines akin to the "fishbone" ones.

Figure 6: Bible of Alcuin.

2.2 Rotating the viewing direction

2.2.1 Mathematical interpretation

In the classical conic perspective, vanishing points positions are bound to the viewing direction. The projection formulas allow us to explicit this position. A conic perspective is defined by a projection center O, a viewing direction, and a screen (a plane orthogonal to the viewing direction) at a distance d from O. These elements define the viewing coordinate system whose origin is O, the z-axis is parallel to the viewing direction—thus the screen is parallel to the xy-direction— and the x-axis is parallel to a given direction of the "screen" (typically the longer one). As this coordinate system directly comes from the world coordinate system by a rotation it is possible to consider that *it is* the world coordinate system. As stated in [FD82], in this coordinate system, the projection matrix is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/d \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (1)$$

Given a direction V of coordinates $(a, b, c, 0)$, and a point M_0 of coordinates $(x_0, y_0, z_0, 0)$,

Figure 4: Urbinate perspective (Berlin panel).

Figure 7: Two vanishing axes.

the points of coordinates $(x_0 + \lambda a, y_0 + \lambda b, z_0 + \lambda c, 0)$ go over the straight line passing through M_0 and parallel to V as λ varies. This point projects itself on the screen point:

$$\left(d\frac{x_0 + \lambda a}{z_0 + \lambda c}, d\frac{y_0 + \lambda b}{z_0 + \lambda c} \right) \tag{2}$$

and thus all the straight lines parallel to V converge to the vanishing point $(da/c, db/c)$.

Reciprocally, given a point (u, v) on the screen, the straight lines parallel to the direction $(u, v, 1)$ converge to this point.

If the viewing direction was turned around the x-axis of the world coordinate system, the resulting image would be the same as if the points were turned in the world by a rotation to the opposite direction before being projected. So the projection is given by the product of the previous matrix and a rotation matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1/d \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{3}$$

Using matrix 1, the point of coordinates (x, y, z) projects itself on:

$$\left(d\frac{x}{y\sin\theta + z\cos\theta}, d\frac{y\cos\theta - z\sin\theta}{y\sin\theta + z\cos\theta} \right) \tag{4}$$

Replacing the coordinates (x, y, z) by those of a point on the straight line parallel to V leads to the coordinates of the projected point:

$$\left(d\frac{x_0 + \lambda a}{(y_0 + \lambda b)\sin\theta + (z_0 + \lambda c)\cos\theta}, \right.$$
$$\left. d\frac{(y_0 + \lambda b)\cos\theta - (z_0 + \lambda c)\sin\theta}{(y_0 + \lambda b)\sin\theta + (z_0 + \lambda c)\cos\theta} \right) \tag{5}$$

As λ increases indefinetely, this point tends to:

$$\left(d\frac{a}{b\sin\theta + c\cos\theta}, d\frac{b\cos\theta - c\sin\theta}{b\sin\theta + c\cos\theta} \right) \tag{6}$$

Hence, the straight lines parallel to the z-axis converge to the point:

$$(0, -d\tan\theta) \tag{7}$$

The vanishing points defined by this formula are vertically aligned on the y-axis. In the general case, the viewing direction should have been rotated around any axis and the vanishing points might have been anywhere in the image.

2.2.2 *Application*

Given an image (a painting), the viewing point is unknown, but several vanishing points may be found for the $(0, 0, 1)$ direction—which is the direction of the parallel beams or one of the direction of the tiles. A viewing direction corresponds to each of these vanishing points.

If the function that associates a vanishing point to each particular straight line parallel to the z-axis was known, it would be possible to compute the rotation of the viewing direction for each point of the world.

In fact, given such a point, the straight line parallel to the z-axis passing through it vanishes to a point given by the function, and the angle of rotation around the x-axis is found using equation 7.

These functions must be interpolated with very few data, looking at picture 5 leads us to suppose that the vertical position of the vanishing point is a linear function of x—but it is very difficult to induce a function for picture 6 because it is not even monotonous! Images on figure 8 are computed with $-d \tan\theta = kx$ (respectively, $k = 0.01$ and $k = -0.01$).

2.3 Moving the eye

[Dam87] explains that the Brunelleschi's experience demonstrates that the "vanishing point" coincides with the "point of view".

> ≪Le lieu où doit se placer le spectateur, ce lieu unique d'où la peinture demanderait à être vue pour produire son effet, trouve son répondant dans le tableau, et que ce soit dans le champ de celui-ci qu'il se laisse d'abord repérer, par *réflexion*. Ce que démontre en effet l'expérience de Brunelleschi, c'est que le point que nous nommons aujourd'hui "point de vue", ce point se trouve coïncider, projectivement parlant, avec le point dit "de fuite".≫[2]

Looking at fishbone perspective paintings, a centered vertical line may be guessed. In order to make this line appear, the eye (the projection center) is translated along the y-axis. The extent of the translation depends on the x position of the point in the world. It should be noticed that the screen doesn't move. Moreover, no care is taken about the vanishing point position or the deformation of the planes parallel to the "screen".

2.3.1 *Mathematical formulation*

If the projection center E is moved to the point (x_0, y_0, z_0), the projection of the point M of coordinates (x, y, z) is the intersection of the "screen" and the straight line passing through E and M; its coordinates are:

$$\left(x_0 + \frac{d - z_0}{z - z_0}(x - x_0), y_0 + \frac{d - z_0}{z - z_0}(y - y_0)\right) \tag{8}$$

Only translation along the y-axis were used, so equation 8 is simplified with $x_0 = 0$ and $z_0 = 0$:

$$\left(\frac{d}{z}x, y_0 + \frac{d}{z}(y - y_0)\right) \tag{9}$$

2.3.2 *Applications*

y_0 is defined as a function of X in the image plane[3]. With such a projection, horizontal straight lines parallel to the screen are drawn as the graph of the translation extent as a function of X. Indeed, such a line is parametrized in λ as (λ, y_1, z_1) and the projection is:

$$\left(\frac{d}{z_1}\lambda, y_0 + \frac{d}{z_1}(y_1 - y_0)\right) \tag{10}$$

which can be expressed as:

$$(X, \alpha y_0(X) + \beta) \tag{11}$$

where α and β are constant.

Hyperbolic perspective The closer the points are to the central vertical line of the screen, the larger the extent of translation. Our first idea was to use a hyperbolic function: the translation extent is proportional to the inverse of the absolute value of X. The resulting images (Figure 9(a) and (b)) show a suction in the middle of the screen.

[2] [Dam87], page 117.

[3] X is the projected point abscissa dx/z

Figure 8: Rotating the viewing direction.

Figure 9: Hyperbolic perspective. (a)Particle method. (b)Ray tracing method.

Figure 10: Different chevron perspectives.

Chevron perspective Reciprocally, the extent of translation may be proportional to the absolute value of X. The "fishbone" effect comes from the absolute value shape (Figure 10).

2.4 Observations on Cubist paintings

Cubist paintings show objects from different points of view. The principle is that the painter turns around the subject; so his eyes describe a circle centered on the subject. Moving the eye around the scene is based on this idea, but it is also necessary to change the screen shape and to replace the plane by an half-cylinder or a cylinder. The rotation angle may be defined in a lot of different ways: we used two of them, first, it may be a function of the distance of the world point (see figure 11(a)), another possibility is to compute each line of the image from a different angle (see figure 11(b)).

2.4.1 Circular screen

The *aim* notion is an usual way to define the viewing direction: in fact, the *aim* point lies on the z-axis. Reciprocally, to give this point and the projection center (the *eye*) defines the z-axis. This notion is used here in another sense as a point on the rotation axis. In fact, this axis is parallel to the x-axis. The aim coordinate system is defined here: it takes its origin at the aim (A) and shares its axes with the world coordinate system.

In the aim coordinate system, the points on the x-axis centered cylinder of radius $(z_A - d)$ verify the equation:

$$y^2 + z^2 = (z_A - d)^2 \tag{12}$$

A point M of coordinates (x_m, y_m, z_m) is projected on this cylinder by finding the intersection of the cylinder and line ME. E is deduced from the origin of the world coordinate system by the rotation of angle ω around the x-axis (see figure 12). A point on the straight line EM is defined by $M + \lambda\vec{ME}$ whose coordinates are:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_m \\ y_m \\ z_m - z_A \end{pmatrix} + \lambda \begin{pmatrix} x_m \\ y_m - z_A \sin\omega \\ z_m - z_A + z_A \cos\omega \end{pmatrix} \tag{13}$$

The intersection P of the cylinder and this line is found by solving the second degree equation:

$$(z_A - d)^2 = [y_m + \lambda(y_m - z_A \sin\omega)]^2 + [(z_m - z_A) + \lambda(z_m - z_A + z_A \cos\omega)]^2 \tag{14}$$

If this equation has two solutions, we choose the one with $\lambda > 0$ because this gives the closest point to the eye. Point P lies on the cylinder, which must be unrolled and this operation is not detailed here.

3 IMPLEMENTATION

The image synthesis system, called ILLUMINES, used at the "Ecole des Mines de Saint-Etienne" [BP90], is based on a CSG modelization. A given scene can be rendered by several hidden surfaces (or hidden lines) algorithms: typically, z-buffers, a ray tracer, scan-line algorithms are available.

The new "perspectives" that were introduced in this paper, are not compatible with some of the most classical and fundamental hypotheses of image synthesis.

With the conic perspective or with parallel projections, straight lines project themselves on straight lines. With the new "perspectives", this is no longer true: for instance, the graph of the eye translation function is the projection of straight lines parallel to the x-axis (cf. 2.3.2). So, the classical hidden surface elimination algorithms cannot be used, because both the clipping step and the filling one work on polygonal data in the image space [FD82].

To circumvent this problem, a z-buffer method was used with particles. A particle is simply a coloured point as in [Ree83]. In fact, each primitive object is decomposed in a fixed (large) number, say one thousand, of particles in the object space. Each point is projected by itself: the formulas (such as formula 4) can be evaluated once x, y, and z are known, whereas it should be difficult (and unusable in the algorithm) to find the projection of a polygon. Computations with this method

Figure 11: Rotating around the scene. (a)Particle method. (b)Ray tracing method.

Figure 12: Circular screen notion.

are quite long, because none of the coherences used in classical z-buffer can be used. Lastly, boolean operations are not currently implemented, but the method described in [RR86] could be used (with even longer computation times!).

Moreover, the light doesn't move on straight line with some of the new "perspectives": this means that the world points which project themselves on a given pixel are not aligned. For instance, equation 4 gives the projected point coordinates. If these coordinates are bound to constant values, two transcendental surface implicit functions are found, whose conjonction gives an intersection curve. In such cases, the ray-tracing algorithm is unusable.

Ray-tracing can only be used when the sole modification is the eye move, and when this move only depends on the pixel position. For instance, with circular perspective, the angle of rotation must be a function of Y position of the line in the screen and must not be a function of the z position of the object.

4 CONCLUSION

Our purpose was to interpret and to approach some of the different representation styles: before the Quattrocento and after the XIX$^{\text{th}}$ century, Cubism. As these representation styles were not based on rigorous mathematics, we tried to extract some formal concepts from pictures. The results that could be obtained were not supposed to exactly reproduce the pictures. But the images seem to be interesting for their "special effects" and for artistic production.

5 ACKNOWLEDGMENTS

Marc Roelens extended his ray tracer to render the images with our special perspectives, he should be acknowledged here for his work.

REFERENCES

[Alb35] Leone Battista Alberti. *De pictura.* Latin Manuscript, 1435.

[BP90] M. Beigbeder and B. Péroche. Un système de synthèse d'images 3d: Illumines. Technical Report 90.2, Ecole des Mines de Saint-Etienne, 1990.

[Dam87] Hubert Damisch. *L'origine de la perspective*. Flammarion, 1987.

[FD82] J. D. Foley and A. Van Dam. *Fundamentals of Interactive Computer Graphics*. Addison-Wesley, 1982.

[Neu67] Beaumont Neuhall. *L'histoire de la photographie depuis 1839 et jusqu'à nos jours*. Belier-Prisma, 1967.

[Pan87] Erwin Panovski. *La perspective comme forme symbolique*. Ed. Minuit, 1987.

[Ree83] W. T. Reeves. Particle systems—a technique for modeling a class of fuzzy objects. *acm Transactions On Graphics*, 2(2), April 1983.

[RR86] J. R. Rossignac and A. A. G. Requicha. Depth-buffering techniques for constructive solid geometry. *IEEE CG&A*, pages 29–39, September 1986.

[Rub90] William Rubin. *Picasso et Braque. L'invention du cubisme*. Flammarion, 1990.

LIST OF FIGURES