

Where Perspective Comes From
Chris Brown
September 2016

Available at <http://hdl.handle.net/1802//31142> ,
<http://www.cs.rochester.edu/u/brown/bio/hobby.html>

Preface

This little auto-tutorial may be useful to others, but was written for me to figure things out. There are many books on artistic perspective, projective geometry, etc, to which I refer you , but to which I did not refer. Critiques and suggestions welcome.

Here I discuss *linear perspective*, one of many conventions for indicating the depth of three-dimensional objects in a two-dimensional image. Other conventions are geometric, like “distant objects are higher” (Fig. 0), and non-geometric: aerial perspective says that “objects get hazier and their color more blue-grey with distance.” Linear perspective is geometric: it shrinks the image of distant objects, which has implication for imaging parallel lines in the scene (as in architecture, roads, rows of trees, etc.),



Fig. 0. Distance and height convention, often seen in Japanese and Chinese painting.

Section 1 describes *point projection* and shows how it produces perspective effects. Section 2 defines common artistic perspective terms. Section 3 is my attempt at a practical guide to using what we learn in 1 and 2. Section 4 mostly relates point projection to eyes, cameras, and art. Section 5 is summary bullets; Sections 6, 7 and 8 are appendices.

1. Point Projection



Fig. 1. By A. Dürer: The pictured artist and vertical grid together implement point projection as he copies the image seen through the grid.

What is (linear) perspective? My informal definition is *the effect produced by depth (distance) in an image rendered in point projection*. Point projection does one primary thing: it reduces the size of an object's image with distance, which can produce changes in angles between world lines. For instance, parallel world lines may not be parallel in the image.

Point projection (Fig. 1) creates perspective effects as it produces a flat (two-dimensional) image from an exterior three-dimensional world. It is a very serviceable description of how a modern camera works. It's powerful enough for realistic computer graphics, and it is our basis for understanding linear perspective. It's simply geometry, so is not (directly) about artistic style but does provide some important artistic choices and constraints on lines and forms; for instance, linear perspective.

1.1 The Pinhole Camera and Point Projection

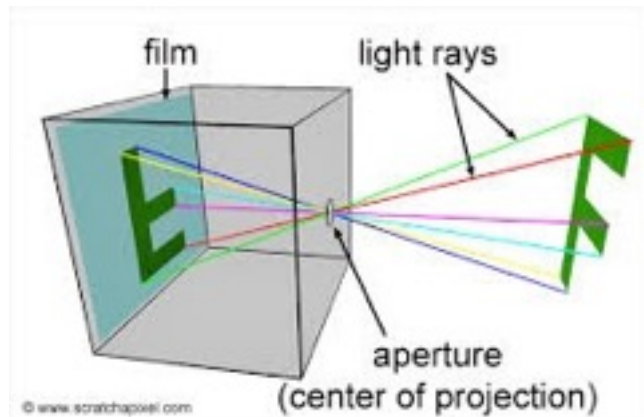


Fig. 2: The Pinhole Camera. Classic portrait cameras reverse the image too.

Point projection is a simple matter of the shapes (geometry) in the scene and (geometrical) optics, which says that light travels in straight lines (rays).

Fig. 1 shows one practical point projection system: another is the *pinhole camera*, which makes an image by receiving light rays from the world through one small hole, (ideally a point, the *point of projection*). Tracing the rays, we see the clever idea (Fig. 2): the hole projects an image on the “film”. A modern camera lens creates the same projection while letting in more light. Thus point projection is a *camera model* (but see Section 4).

The pinhole image is upside-down and backward. Let’s say the *image* is produced on the *image plane* (usually a flat sensor or “film”, as Fig. 2. calls it). In this paper the camera is always aimed along a *gaze direction or gaze axis* pointing “straight out of the lens” perpendicular to the image plane.

Finally, *point projection* simply puts the image plane in front of the point of projection (pinhole), not behind it. The pinhole is now called a *viewpoint* (Fig. 1 shows why). The viewpoint is now on the gaze axis some distance behind the image (see Fig. 3. below). This trick changes nothing essential in the analysis of how point projection works, but it does remove the pinhole camera’s annoying direction reversals (seen in Fig. 2.). In Fig. 1, the artist’s (copyist’s) eye is the viewpoint and point of projection and the image is projected “on” (really “through”) the vertical grid.

1.2 The Side-view Camera and Point Projection Equation

Fig. 3. is a side-view of the point projection process. In a useful simplification, the camera (“eye symbol”) and the gaze axis are on the floor, in the *ground plane*. The “eye” on the left of Fig. 3 is the viewpoint, the line I is the image “plane” (seen edge-on as a line), and the distance f is called the *focal length* of the projection: it’s what changes when zooming a zoom lens. The edge-on “image” is one-dimensional (a line, not a rectangle), with an object’s world height h and the focal length determining its image height h' . Distant objects are imaged smaller than closer ones of the same height (Fig 3 top), and a shorter focal length accentuates the shrinkage (bottom).

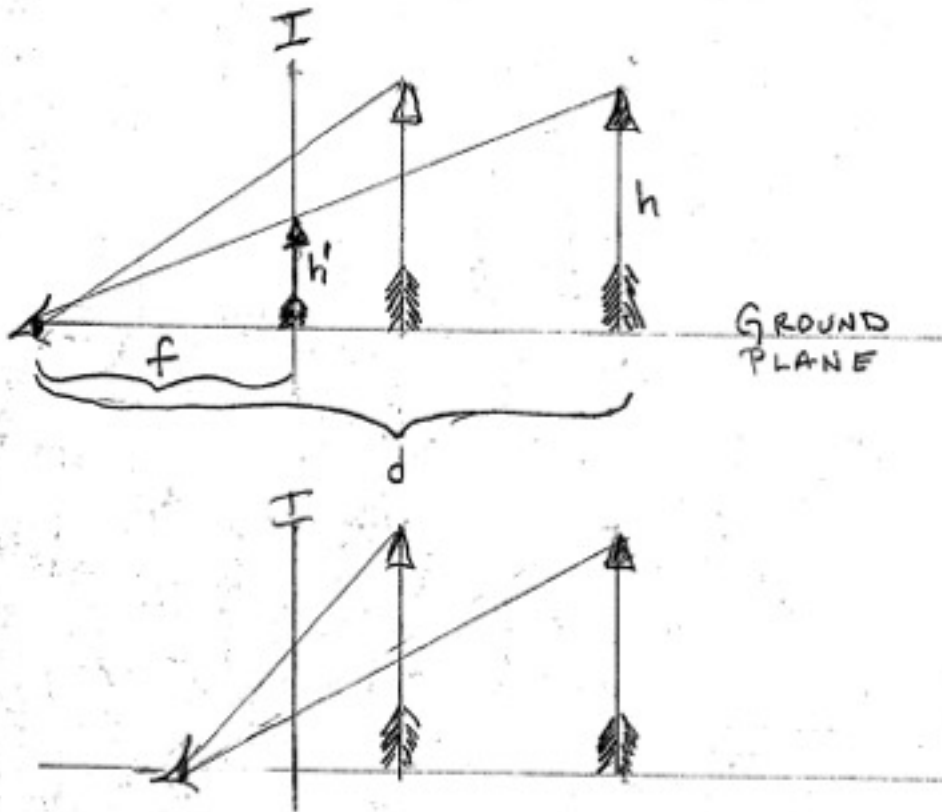


Fig. 3. Point projection simplified to one-dimensional imaging. Here the gaze axis lies along the “floor” (ground plane). Top: As object distance d grows, the image size h' shrinks. Bottom: Shorter focal length f exaggerates perspective shrinking.

MATH TRIGGER WARNING! OPTIONAL. If you did trigonometry in middle school you recall, or it's easy to believe, that *similar triangles* called A and B, with B's three angles all the same as A's, "look alike" because the angles are the same and also the size of A (its side lengths) is a scaled version of the size of B. In particular: in (Fig 3 top), the triangle with sides f and h' is similar to the one with sides d and h . The arrows in the world both have height h , and h' is the height of the far (rightmost) arrow's image. The definition of similarity means that in these triangles

$h'/f = h/d$; that is, the smaller one's side lengths are scaled versions of the larger's. Thus $h' = (fh)/d$, (multiplying through by the focal length f).

This projection math is the same for the "width," or w direction, as for height, or h direction, so for any point (w, h, d) in three-space, its image is at some (two-dimensional) image point (w', h') given by $(fw/d, fh/d)$. That gives the complete definition of point projection:

Point Projection Equation:

With focal length f ,
the image of a 3-d world point (w, h, d) is at $(w', h') = (\frac{fw}{d}, \frac{fh}{d})$ in the image.

This simple relationship completely describes point projection and shows (for instance) that moving the object away from the viewpoint (increasing d) decreases its image size since we're dividing its actual height and width by a bigger value of distance. **END OF MATH.**

1.3. Summary and Related Facts

- * Point projection is an imaging technique that reduces the size of distant objects according to a formula easily derived from the simple projection geometry. In a deep sense, any image from point projection is always "in perspective" everywhere and consistently.
- * It is not obvious, but point projection of any straight line in three-dimensions is a straight line in the image. We also do not prove the crucial fact: *a set of parallel world lines, not parallel to the image plane, converge in a single "vanishing point" in the image.* You can convince yourself of this just using the perspective equation, or graphically for vertical or horizontal lines using the one-dimensional simplification of Fig. 3.
- * Section 2 categorizes four increasingly general cases of the point projection of parallel lines, with the aim of relating point projection to the art-theoretic notion of "N point perspective" for $N = 0, 1, 2, 3$.
- * If focal length f is very short (a "wide angle" lens), then it has a wide field of view, and also small changes in d cause big changes in h' (image height). This effect can, e.g., increase the size of a nose in a close-up frontal portrait. Portrait photographers use semi-telephoto (long f) lenses to avoid that (among other reasons).

- * It's useful (say, to Hollywood) that any point-projection 'camera' is totally described by only a dozen or fewer numbers, and that there are simple mathematical ways to move and aim the model camera, so the computational basis of graphics (producing an image from a 3-D model) is very simple.

2. Zero- to Three-Point Perspective and the Horizon

So now we know about point projection, but what are and why do we care about the traditional art-class ideas of “one-, two-, and three-point perspective”, and “the horizon?” Let's see what these concepts are: are they descriptive, proscriptive, or what?

2.1 Zero-point and One-point Perspective

The side-view of our point-projection camera (Fig. 3) makes it clear that as f increases, the image height of an object approaches its world height. If the focal length is infinite the viewpoint is infinitely far to the left and the lines of projection are all parallel to the gaze axis; projection causes no perspective shrinkage. This case of "parallel projection" is called **orthographic projection**. It is approximated by telephoto (large f) lenses, like the center-field camera of baseball TV coverage: the pitcher, batter, and fans in the stands look the same size (Fig. 4).



Fig. 4. Telephoto (long focal length) lenses minimize shrinkage of size with distance, and approximate orthographic projection.

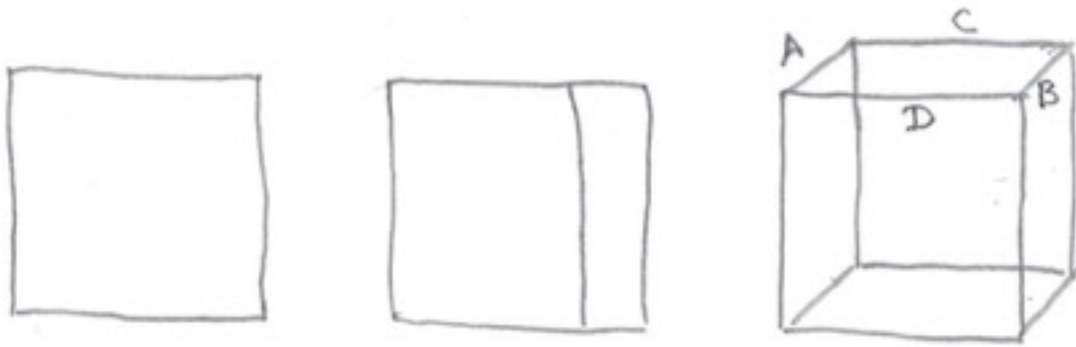


Fig. 5. Orthographic projection of a cube. Parallel world lines are parallel in the image. ABCD is a parallelogram. There are no “vanishing points.”

Orthographic projection preserves parallelism since there is no perspective shrinkage (see below), so in projection the cube’s faces are parallelograms (since lengths can be foreshortened but not perspective shrunk in depth) (Fig 5) . Orthographic projection is used in mechanical and some architectural drawing.

I’m glad to introduce (at last) a perspective concept that you may have heard of in art (or art history) classes: Fig. 6. shows **one point perspective**. It shows three point-projection images of a cube whose front and rear faces are parallel to the image plane. Its top and bottom faces are parallel to the ground plane. The front and back faces are at two different but constant depths and are imaged as squares that differ in size. Faces whose depth varies are imaged as trapezoids: lines A, B, C, D have lost their worldly parallelism and meet in a *vanishing point* in the image. If this cube’s orientation with respect to the image plane is fixed, any point-projection image from any viewpoint will be in one-point perspective.

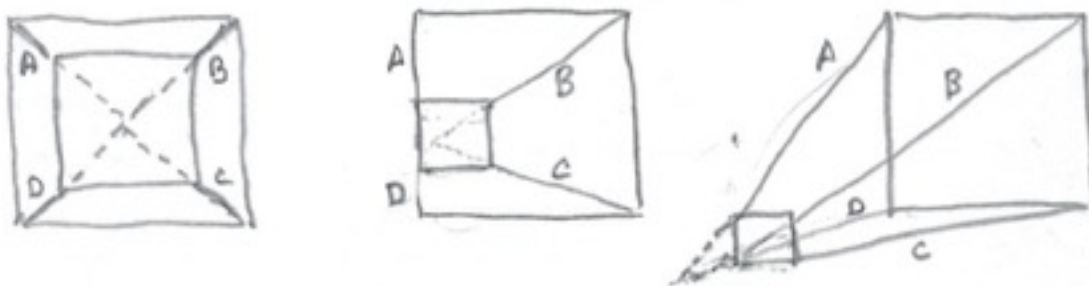


Fig. 6. One-point perspective. Left: gaze axis through face centers, moderate perspective shrinkage. Middle: viewpoint moves left and up so gaze axis is in a face-plane of the world cube; a shorter focal length causes more shrinkage. Right: Even shorter focal length, viewpoint now left and below ground plane. Dashed lines extend image lines to the single vanishing point.

Generally, we have one-point perspective if *there is a single set of lines in the scene all parallel to each other but not parallel to the image plane*. In Fig. 6, those are lines A,B,C,D.

We could call orthographic projection *zero-point perspective* since it has no vanishing point.

Foreshortening: Foreshortening is the “shrinking” effect of viewing an object not head-on in 3-D but at a slant. For example, the left and right faces of the cube’s image in (Fig. 5 middle) are each a foreshortened view of a square face (seen as such in Fig. 5 left). Important: there is no perspective effect in Fig. 5. In general, point projection with finite focal length produces both foreshortening and perspective shrinking. However, it is sometimes useful to distinguish the two phenomena (for example, in some physics applications foreshortening is important but perspective is not). Often the foreshortening effect dominates the perspective effect: adding a little perspective distortion to Fig. 5 middle would be a small effect compared to the foreshortening of the faces. Section 8 has more on foreshortening.

Beyond Cubes (quasi-optional). Let’s stretch a bit. First let’s imagine an image with two cubes, X and Z. X is like the cubes in Fig. 6, rigid with standard right-angle corners. Z is different. Its front and rear faces are rigid but its equivalent of lines A - D, while rigid with constant length, are not rigidly attached at the cube’s corners: they can swivel around the junction to point in any direction. If the front and back faces had weight then gravity would pull the back face down to hang below the front face with edges A-D hanging vertically down and still connecting front and back. So now imagine we pick up the back face, move it so that seen from the front it is to the left of the front face and below it. We now have front and back faces parallel but we no longer have a cube or even a rectangular parallelepiped. We have a parallelepiped skewed in two directions, in which the edges A-D are parallel but not “headed straight back” perpendicular to the image, as are those of Fig. 6.

If we use what we learned and image Z as we did X, we expect its lines A,B,C,D to have a vanishing point, and one that is different from the vanishing point of X’s lines A—D. X and Z are two independent one-point perspective situations, and can co-exist within the same image. In this paper I tend to call this a two-point perspective situation since there are two vanishing points, but you could say each relevant local perspective-rendered region (around X and around Z) is in one-point perspective. **End Beyond Cubes.**

Summary: One-point perspective (in a structure or region, say) means that there is only one set of receding parallel lines, which have a single vanishing point. Point projection makes the image of a cube’s far face smaller, pulling the edge images together to a common vanishing point in the image plane (possibly out in the finite image we’re drawing). There are thus two related perspective effects: size and angle changes.

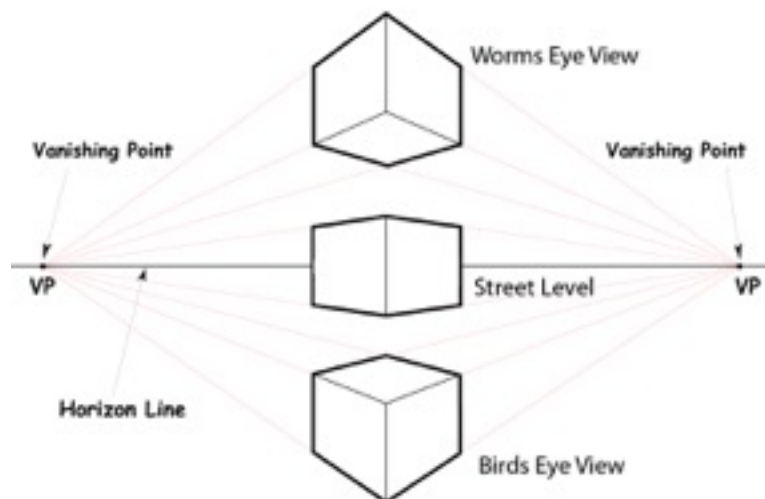


Fig. 7. Image of three cubes at different heights: middle cube's height centered on the gaze axis. We see two vanishing points, hence *two-point perspective*, and a horizon (Section 2.4).

2.2 Two-point and Three-point Perspective

We have seen that **two** sets of parallel lines, neither set parallel with the image plane, each set pointing in a different direction, create two vanishing points in point projection. Our hypothetical situation above had to “stretch” to create this situation — to keep using our one-point projection definition we needed a skewed object. Much more natural is just generalizing our imaging geometry to realistic situations by twisting our previous cube (Fig. 6) around a vertical axis so it has no faces parallel to the image plane (Fig. 7 shows three such cubes).

Our twisted cube now has two sets of parallel world lines changing in depth, imaging to lines that meet in two different vanishing points if extended. The vertical world lines remain parallel to the image plane so stay parallel in the image. Two vanishing points means *two-point perspective*.

The vertical lines in Fig. 7 are parallel to the image plane since we're gazing “straight ahead”. So we can make another vanishing point (get *three-point perspective*) by twisting the cubes again.

First, it's useful to remember that the image changes if either the object or camera moves. We rotated the cube of Fig. 6 to get Fig. 7, but we could have created the same image by moving the camera (rotating the viewpoint) in the opposite way (thus preserving their relative geometry). Likewise here, we can tilt a Fig. 7 cube backwards away from us, or we can “gawk” and rotate our gaze direction upwards (as when we look up at a tall building), which rotates our image plane to lean back over our viewpoint. Either way the used-to-be vertical lines in the world cube are now pointing “away from” the image plane: they produce three vanishing points. (Fig. 8).

Three-point perspective may often seem like “overkill”, since the world’s vertical direction (of gravity, hence of tree trunks, building columns, human posture, etc.) is parallel to the usually vertical image plane of our eyes and cameras.

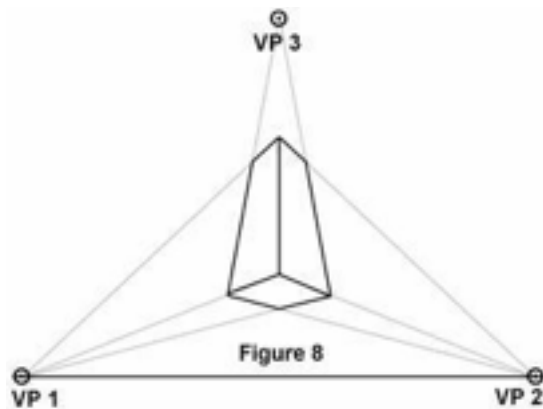


Fig. 8. Three-point perspective and horizon.

2.3 Important: Mixed Perspectives

This little section addresses the “status” of vanishing points and of N-point perspective in general. Practically, they are determined by the scene; they don’t govern the scene, and are not arbitrary artistic choices. We don’t usually just make them up when beginning an artwork.

Pictures of scenes are not necessarily “in” one perspective type (0 —3 point). Don’t forget what is really going on, which never varies: it’s always **point projection**. The perspective type (if any) is dictated by viewpoint and the world geometry of the structure depicted. Example: the scene is a cubical room with a cube floating in it. The viewpoint is fixed, with its image plane parallel to the far and near walls, so the room’s edges appear in **one-point perspective**. BUT the floating cube is rotating continuously, slowly, around all three axes at once: through time it is oriented every which way, like some screen-saver. It coexists with the unvarying one-point perspective of the room lines, and indeed it is within the volume they enclose. BUT as it rotates, its edges usually *all* vary in depth: only by very improbable accident would it align with the room’s walls, say. SO it has no privileged “up” direction, and the only consistent way to deal with it is with **three-point perspective** (Appendix 2).

But WHICH three (vanishing) points? If you wanted to animate its motion (draw it at many instants) would you pick vanishing points for each instant and then draw it? You would NOT, because they are hard to reason about: their motion is not easily derived from (or related to) changes in 3-D line direction (Appendix 2). If you get them wrong they won’t be consistent with the desired cubical shape. What you would do (you must do, I claim) is shoot a movie (literally) from your fixed viewpoint and copy the movie frames, deriving from each frame the correct vanishing points to help you create a consistent and accurate cube drawing. OR, if you are making up the scene, you would use your subjective perspective skills to draw each frame in

sequence so it looks “right” to you. You would use each one to determine vanishing points, which you would then use to correct and perfect your first drawing. You’d *discover* the vanishing points, not dictate them.

2.4 The Horizon

In two-point perspective, both the vanishing points of the edges of a cube *on the ground plane* lie on a horizontal line in the image (Fig. 7.). Now imagine spinning the cube around a vertical axis: as it spins, it creates an infinity of 2-point perspective situations as its vanishing points sweep horizontally. They must lie on that same horizontal line: the *horizon* is the locus of all these vanishing points, indeed of the vanishing points of all sets of parallel lines in the ground plane (Fig. 9 top). A more familiar depiction is (Fig. 9 bottom); here the lines are parallel to the image plane so they have no vanishing point, but the perspective shrinking of their spacing creates the same horizon).

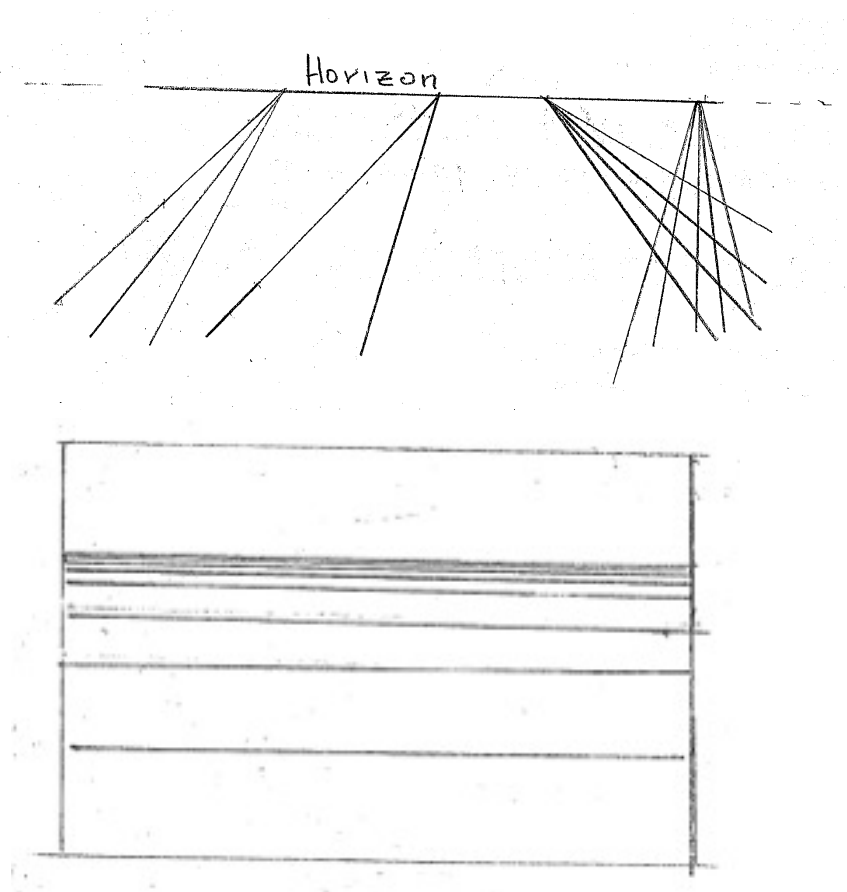


Fig. 9. Top: Horizon contains all vanishing points of parallel lines in the ground plane. Bottom: Lines with no vanishing point can determine a horizon.

Fig. 10 is a side-view of what's happening in (Fig. 9 bottom).

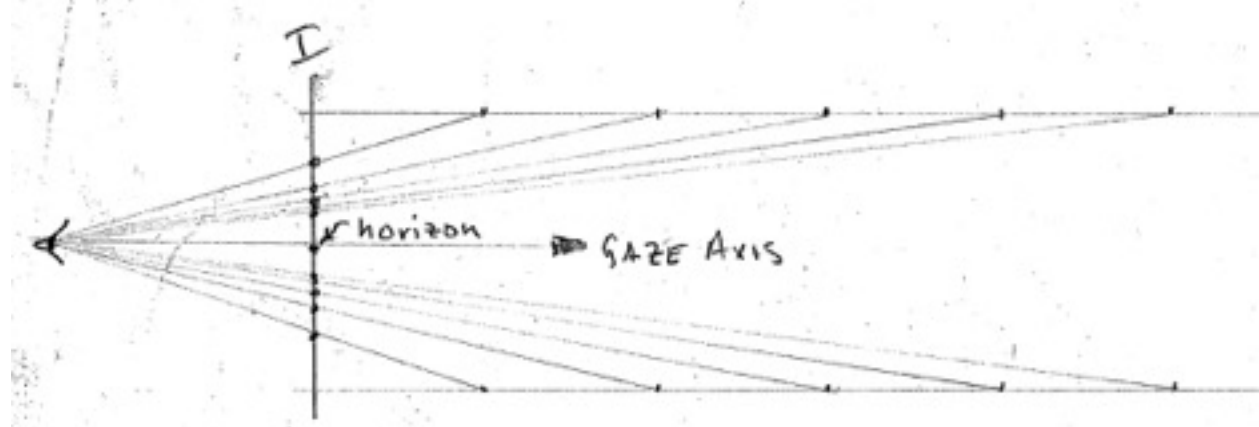


Fig. 10. The side-view camera looking right, over a horizontal ground plane like the one in (Fig. 9 bottom), and under a ceiling plane similarly ruled: from the side we see the ruling lines end-on, as points. Likewise the edge-on image (imagine (Fig. 9 bottom) seen edge-on) is the dots on the I plane, and the gaze axis intersects the horizon line (seen end-on as another point).

The horizon is an image phenomenon, it's not out in the world. It gives us a continuous, infinite set of vanishing points, which is cool, but maybe not the most useful way to think of it. All planes parallel to the ground plane share its horizon (Fig. 10.). Therefore if objects in the world have some maximum height, and they are attached to one of these planes (sitting on the ground or hanging from the ceiling, say), then *the horizon is where they all shrink to nothing*. That's useful, and doesn't refer to (need no stinkin') parallel lines.

3. Methods and Monsters

3.1 Methods

How to use all these ideas practically? This section seems to me the logical consequence of what I've learned about perspective (by writing this) but I've never read a serious book on perspective, nor heard the following "method" in a lecture, nor used it in practice. This section is my effort to convert "my understanding" to practice. Appendix 1 is a related topic. As always, suggestions and objections more than welcome.

1. We assume you want to use point projection to create your image. Find or imagine your 3-D scene, put your real or imaginary self in it at your choice of viewpoint. In a real scene your imaging "focal length" is determined for you, by scene geometry and your eyes. If you are making things up, you can exaggerate the effect of depth or attenuate it by imagining your desired effect, thus making up your own focal length.
2. Next, regard your scene to see if there are important sets of parallel lines in the world, either explicit or implicit (like the imaginary line connecting the tops of fenceposts marching off in the distance). One source of parallel lines is the inside or outside corners of rooms or buildings. For 3-point perspective you need at least two parallel lines each for "width, height, depth": two corners are enough if they have the same angles (often they're all right

angles but not necessarily) and do not share an edge line (that would only be five lines total). If you can't find enough real or imaginary parallel lines or they aren't helping you organize your approach, you can use viewing or measuring aids as in Fig. 1 to duplicate the effects of point projection: you “don't need no stinkin' vanishing points”.

3. Using the real or imagined parallel lines, locate the vanishing points in the image you see or are creating. They are emergent phenomena, which is a polite way of saying they are just symptoms; they are hard to predict but are determined for you by the imaging geometry (See Section 3.2, Appendix 2).
4. Remember Section 2.3. Different parts of the scene may (well) call for different types of perspective, which you can recognize using Sections 2.1 and 2.2. E.g. from a police helicopter hovering over a suburban development you see a typical dead-end traffic circle with five cube-like bungalows ringing the circle. There may well be enough corners and lines to put each of those houses, and the ones lined up in parallel along the straight drive, into 2- or even 3-point perspective locally, not affecting the others.
5. Now you can draw (helper or imaginary) lines through the vanishing points to reflect (and guide) the placement of objects and lines in the scene.
6. If scene objects are “normally” related by function, design, or physics (unlike the room and cube of Section 2.3) their images may be related and may share perspective characteristics: an object in perspective gives guidance on how a related object should be rendered consistently with it (Section 6, Fig. 15).
7. When you run out of explicit aid from linear perspective on drawing directions of lines or sizes and shapes of objects, your job is to finish the work following the perspective rules you have chosen or (more basic) point projection. Subjective perspective (drawing what you think or deduce your projection would produce) comes into play.

3.2 Monsters

To re-reiterate: Vanishing points are in the image, not in the world. They result from parallel world lines and imaging geometry. They are not an arbitrary choice like viewpoint, so we should be guided by the real or imagined situation and point projection, rather than thinking “perspective drawing means choosing vanishing points”. Two examples follow.

3.2.1 Bogus One-point Perspective

You see an “obviously” two-point perspective situation but decide to use one-point perspective for your personal style or to express solidarity with primitivity, say. You can “do that:” just let one of the vanishing points go to infinity (way out of the image) —it ceases to affect lines in that direction, which now stay parallel in the image. You get a cube like Fig. 11, which is indeed an image but not one from point projection since two focal lengths created it.

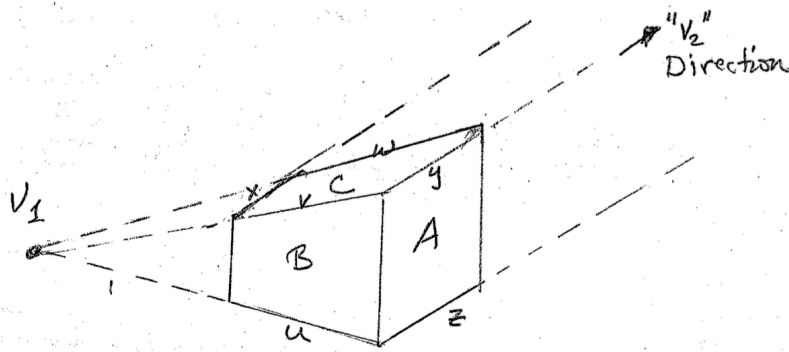


Fig. 11: A cube with one vanishing point $v1$ in the image and one ($v2$) at infinity gives an inconsistent “one-point perspective.” It cannot be made by point projection since the focal length would have to be both finite and infinite to produce $v1$ and $v2$.

3.2.2 Free-floating Vanishing Points

If the artist can pick vanishing points at will (he can't, is my point), consider letting them move around and see what happens. (Fig. 12 left) shows a perfectly normal two-point projection with two fine vanishing points, $v1$ somewhere off up left and $v2$ off up right. You decide to let the two vanishing points start moving toward each other. Why not? They're yours, yes? (Well, no, they're actually not — they belong to the viewpoint, focal length, and scene). At some time $v1$ and $v2$ will line up vertically above the scene, which will become the vertical line of (Fig. 12 middle.) As they keep moving, the cube is “pulled through itself” to become some perspective view of the mirror image of the original (Fig. 12 right). “Verrrry interresting...but not funny.” This monstrosity uses the same rules as normal imaging, but now they violate physics: more constraints on vanishing points.

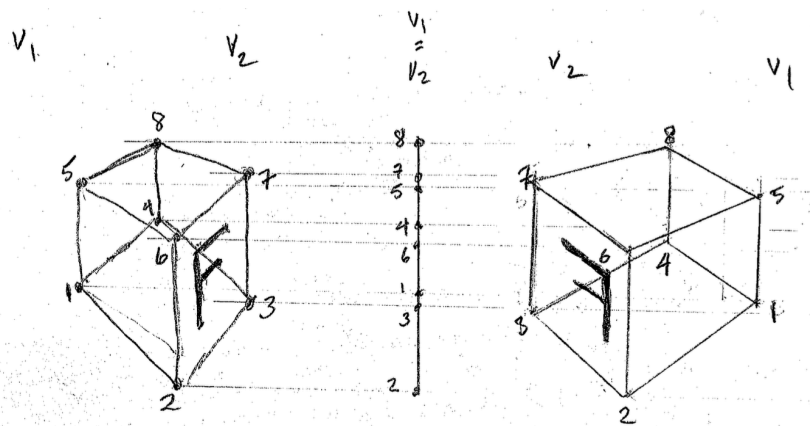


Fig. 12. Two vanishing points approximately exchange places, showing associated images. Left: starting position. Middle: they superimpose or line up vertically. Right: ending position.

4. Perspective, Biology, and Art

4.1 Point Projection, Cameras and Biology

Is the point projection model privileged or “right”? No, it’s a mathematical model and a convention, (each a compelling one). We noticed in Fig. 6 that it is easy to make some pretty violently “distorted-looking” projections with our model camera. Why do they look unnatural? What else goes wrong? Well, point projection is not exactly (well, even remotely) like a real imaging system.

Our mathematical camera has an infinite image plane and all points, lines, and details in the 3-D world are projected on it with 100% precision and perfect resolution. We leave aside the fact that our eyeball has a spherical imaging surface; there are bigger problems than that.

Our eye has a pretty big field of view, (almost 180 degrees) but a standard cameras has only about 30. That is, a camera only images the scene close to the gaze axis. Much worse and more importantly, we humans only “see” things that are pretty exactly **on** our gaze axis. Our retina (“film”) has highly varying resolution, which is best at the *fovea*, a very small area (covering an area in the visual field about the size of a thumbnail held at arm’s length). At even small angles off our foveal gaze-axis, we become legally blind, then worse. This is easy to test. Find a page of large print, force yourself to gaze at its center, and note how the letters become unreadable with distance from your foveal gaze-point. Fig. 13. shows the size of world objects that we see with the same detail if the gaze axis is in the center.

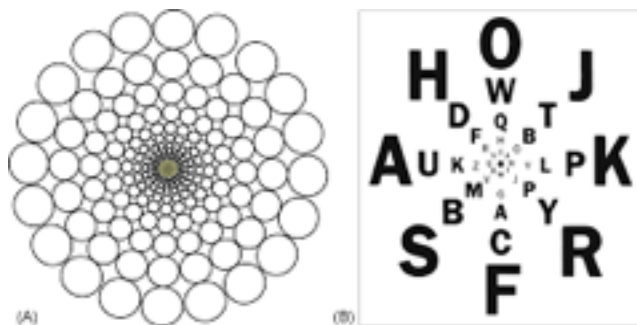


Fig. 13. High resolution in the human fovea and drastically lower “peripheral vision” resolution means that in daily life we quickly shift (rotate) our eyes to point the fovea where we want to see details in the scene.

We all have a miraculous illusion that our visual field is in focus and stable, but our perception is created from a stream of multi-resolution foveal-peripheral scenes as we move our visual attention (and gaze direction), purposefully or reflexively, several times a second. Eye

movements have the effect of moving the image plane by re-aiming the camera, which changes the effects of perspective on different parts of the scene as we naturally (often unconsciously) look around. For drawing, the point projection is static, as from a fixed camera whose gaze angle does not change.

4.2 Problems in the Periphery

4.2.1 The Non-shrinking Squares

Consider the point projection image of an infinite plane of squares parallel to the image plane: the image is a grid of (maybe smaller) squares, off to infinity. They do not shrink with distance from the viewpoint (they're all the same distance from the image plane, which is what point perspective cares about (Section 1.2)). They are not increasingly foreshortened, they are not affected by perspective, they are not squished rectangles, they are squares. The image looks like graph paper. That is counter-intuitive and seems to point to something deeply 'wrong' with the imaging model. That's all true (Section 4.1).

What we humans would do if we wanted to see some square on that face, in the distance off to the right, is "look at it", i.e. change our view axis to put it on our fovea. Boom. The square now shrinks due to the increased distance (to the image plane), is foreshortened due to its small angle to the gaze direction, and itself shows perspective (becomes trapezoidal) since its right edge is now farther away from the image plane than its left. So point projection of off-axis objects is not really wrong (the math is what it is) but does not accord with the way humans (or their cameras) make the image.

But we may not want a true perspective image: if we're photographing the Empire State Building from the sidewalk, gazing up at it makes its size shrink with altitude, and its parallel sides seem to meet off at some vanishing point (Fig. 8.) If we wanted its sides to be parallel in our photograph, we could use a *swing-back camera*, which alters the angle of the image plane to the gaze direction. Using the swing, we could make the image plane parallel with the vertical frontal plane of the building, and its sides would stay parallel in the image.

4.2.2 The Elliptical Sphere.

Another example: have you ever noticed perspective effects for a sphere? Try it! Get a (soccer, tennis, ping-pong) ball: can you find a position in which it is perspectively distorted? I think your answer will be no. If infinite, high-resolution point projection were how our eyes worked, we'd notice that the off-axis sphere is an ellipse (intersection of cone and slanted plane) (Fig 14). I claim we humans just can't do it: if it's off-axis enough to project as an obvious ellipse, it's in our visual periphery, too poorly resolved to get any shape information whatever. In "real life", when we're interested in a sphere, we actively "look at it" (projectively, aim our gaze axis at it; biologically, redirect our eyes to image it on our fovea), and it then projects to, and so looks like, a circle.

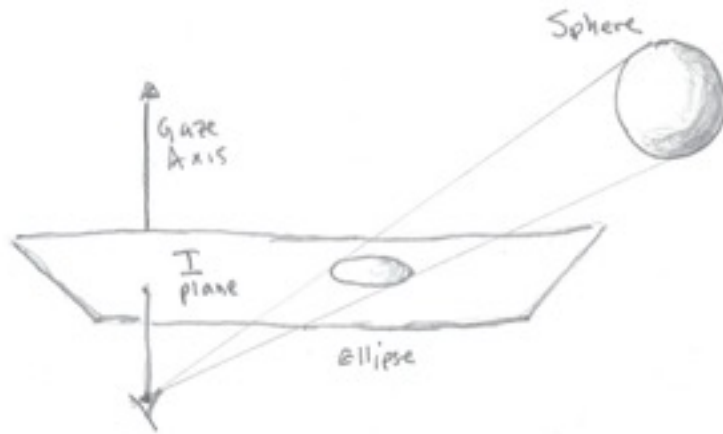


Fig. 14. An off-axis sphere projects to an ellipse in the point-projection model. We humans never see this ellipse because we must direct our gaze axis to the sphere to see its details.

4.3 Opportunities: 0-3 Point Perspective as Constraint and Information

Our real-time world is nothing like a photograph or drawing: it's actively constructed from many viewpoints. But as artists we can pick an instant in time and a single point of view, and decree them to be fixed.

We've seen that we discover vanishing points depending on local evidence, and said that this local imaging geometry may (helpfully) constrain the look of nearby objects, whether they have parallel lines or not. So a local object or environment (room, street, cube, floor) governed by simple perspective may give information on sizes and orientations for other objects in the scene. This effect is at work in "optical illusions" using perspective. Fig. 15 is typical: The perspectively-imaged floor implies a greater depth for the upper sphere, which thus must be larger than the lower one to appear the same size, so it "looks" larger (but it's not).

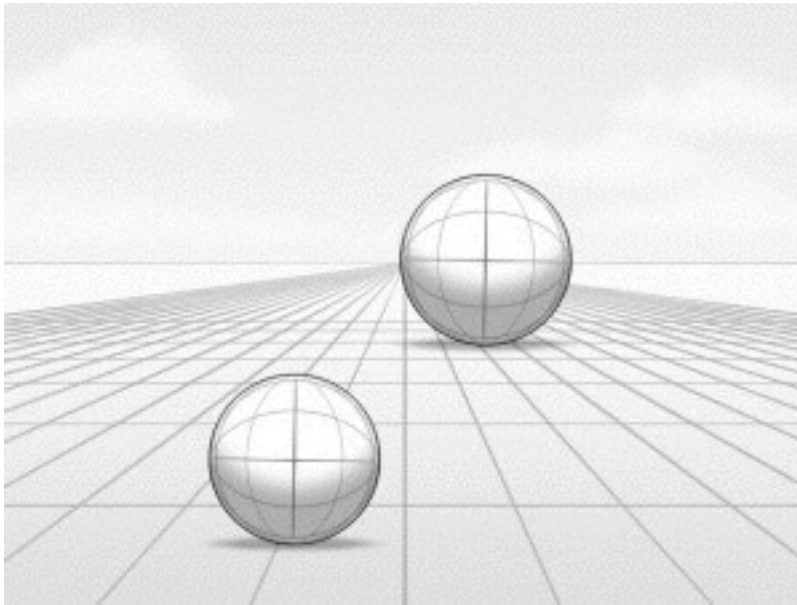


Fig. 15: An “optical illusion” caused by perspective clue

A non-illusory example is the “box-top” scene, in which we glean information about a part of the scene not determined by, but “living in”, a simple perspective situation. Fig 16 is a box in 3-point perspective with a top that is open.

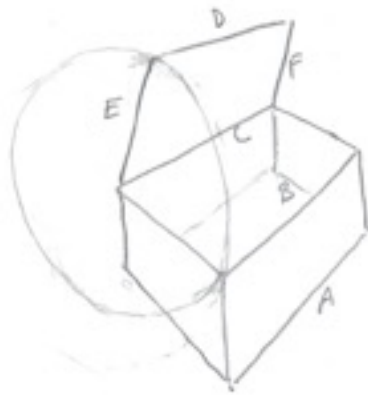


Fig. 16. Four point perspective: how does theory help? (see text).

What can we say about how to draw the top? Lines A,B,C are parallel in the world and share a vanishing point. But line D must also be parallel to them: with top closed, D is the same as B, and rotating around C (parallel to B) maintains the parallelism of D. So we know to draw D so as to share (intersect) the A, B, C vanishing point. Now we're sort of stuck: how to draw lines E and F? We know that the point on the top's near corner rotates through space in a circle with radius of the box's side, which forms a perspectively-distorted ellipse in our image: that may help, since now *subjective perspective* must take over. Subjective perspective is simply "drawing what you think point projection would produce; what looks realistic and undistorted."

We conclude that knowing about the perspective properties of a scene object can help us reason about the image of a related object, even if we cannot say exactly how it must look without more information.

Of course if we have an explicit geometric model of the box, as we would in a 3-D graphics program, we can use the perspective imaging equation from Section 1.2 to calculate where any known world points (like the corners of the top) are imaged. Likewise the rotating box of Section 2.3: it takes about four lines of code to compute where a point (and its image) is, and one or two more to find a vanishing point.

4.5 Beyond Classical Perspective

As I said on page 1, point-projection isn't the only imaging model: consider curved movie screens, fish-eye lenses, etc. Artists go beyond, around, and through restrictions all the time: see Fig. 20. One contemporary example: Brian and Trevor Oakes (at oakesoakes.com) use optical devices, curved canvases, I don't know what-all, to create non-standard but compelling imaging models and resulting images (Fig. 17).

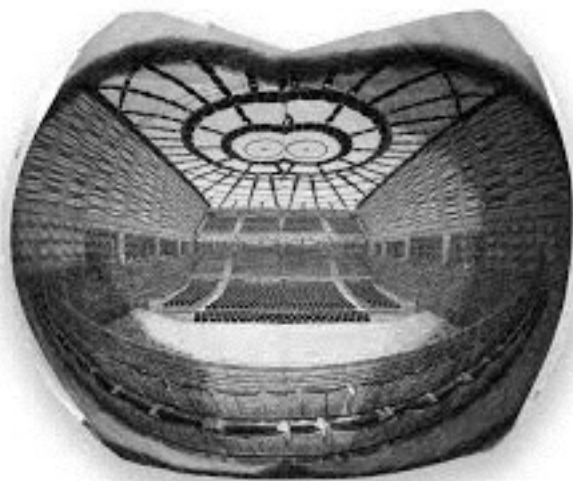


Fig. 17. Concave canvas (I think): one of many images at the Oakes brothers' website.

5 . Summary

- * Point projection is a mathematical model of an optical process that produces images in linear perspective.
- * 0, 1, 2, and 3-point perspective describe what happens to object sizes and (with restrictions) parallel lines and parallelepipedal volumes under point projection.
- * A vanishing point is where a projected set of parallel world lines (extended as necessary) intersect in the image: N-point perspective has N vanishing points.
- * The horizon is the image of an imaginary line in the world where distant objects on the ground plane shrink to nothing if they all have some maximum size in the world.
- * Strictly, 0-3 point perspective applies only to a limited set of world situations, but combined with subjective perspective can help produce more “realistic” (photograph-like) images.
- * It is common that N-point perspectives for different N and with different vanishing points occur in a real scene. The process is governed by point projection, which creates vanishing points from parallel lines as necessary.

6. Appendix 1: Cells in Space

I thought this was a cute idea but now I’m not so sure its useful or even explanatory. Its limiting assumption is that every object in a reasonable-sized volume of world space is governed by the same zero to three vanishing points. I describe this technique initially for one-point perspective, and also as applying to the whole imaged world. Two and three point perspective appear in the last paragraph, and the technique can be applied locally to any volume of world space you wish, like those that have different but coherent sets of parallel lines (say a village of randomly oriented houses). For the volume you want to consider, imagine that all of its space on the far side of the image plane is divided up into a 3-D ‘grid’ of identical cubical cells, each with front and back faces parallel to the image and top and bottom faces to the ground plane, and two facing directly left and right. That is, the edges all line up with the width, height, depth (w, h, d) directions of Section 1.2.

Any layer of square faces (like the front plane of the cell array) parallel to the image plane projects to an image of equal-sized squares in the same vertical and horizontal alignment. This is rather surprising, since we might well imagine that the squares get more foreshortened with “distance” off to left and right, and that the image would be related to Figs. 9 and 10. But in Figs 9 and 10 the image plane is *perpendicular* to the plane of the rulings: our front face of cells is *parallel* to the image plane, which makes all the difference. (An easy way to convince yourself is to use the projection equation; almost as good is to draw the equivalent of Fig. 10 only with a vertical ruled world line to the right of the Image plane — its image points are equally spaced too!

Given this imaginary 3-D cell grid, we may choose any point in the image plane as the vanishing point. All the backwards-pointing cell edges now realign to go through that point. We have thus made images of all the cells: we now know how all those cubes throughout space will look in the image. And we can use any rectangular parallelepiped instead of a cube for a cell, and can translate and resize cells arbitrarily, and produce their images the same way. Besides face shapes, we know how all the right-angled corners of all the cells look.

The cell images are little maps that show how the simple cell and its edges distort under perspective. If a cell surrounded a human head with features, that head's image stays in, and is distorted like, the cell's. We can't read off the appearance of curved surfaces, but with the help of the cells (we can subdivide them for better resolution) we can use *subjective perspective* (what we guess the image should be, using interpolation, extrapolation and artistic judgement, etc.) to help us draw the head.

This actually seems powerful: we can calculate the image of any cell in space for any fixed vanishing point. But there's more: We can rotate the 3-D cell grid to present two faces to the image plane: this leads to a richer perspective world with two vanishing points, which may be more realistic, pleasing or more accurately reflect the desired viewpoint — two point perspective! Vertical depth variation is added with a third vanishing point and we get three-point perspective. At this point we have a vanishing point that affects every cell edge, and adding more vanishing points doesn't make sense. Three-point projection is as far as we can push this idea, but since the idea can be applied to any locale in world space, we can subdivide any world volume this way into locally-coherent 0-3 point perspective cells.

7. Appendix 2: One-point and Three-point Perspective (Section 2.3)

Figs. 18 shows the images of a scene that changes from one- to four-point perspective as a cube rotates in a static room. Fig. 19 shows how the three vanishing points of the cube change as it rotates at constant (but slightly) different rates around the horizontal, vertical, and depth axes.

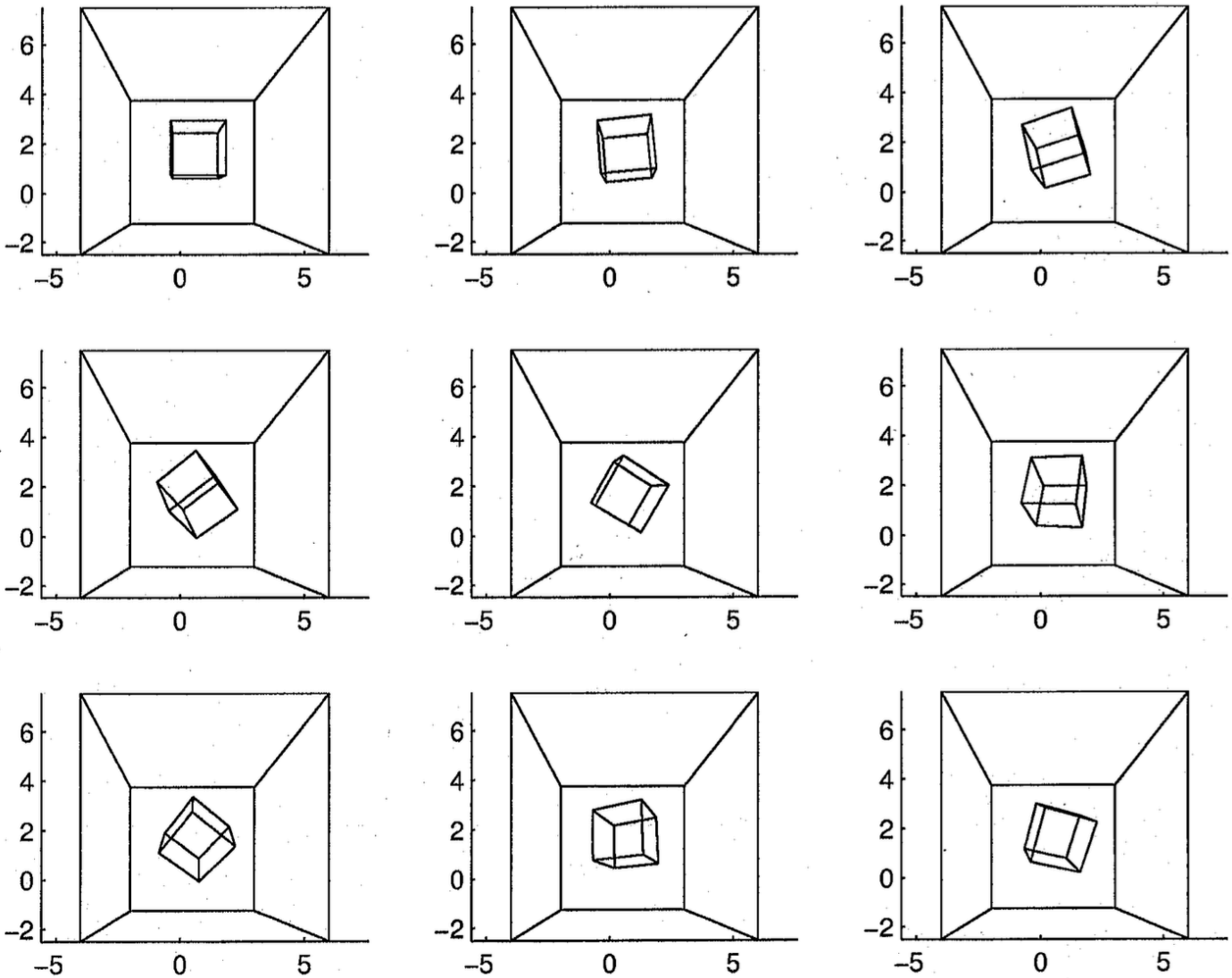


Fig. 18. As described in Section 2.3. Rotating cube in room. The viewpoint gives 1-point perspective for room (and initial cube position). As floating cube rotates, the room stays in 1-point perspective (vanishing point in the viewing direction, or image point $(0,0)$), and the cube displays 3-point perspective caused by point projection.

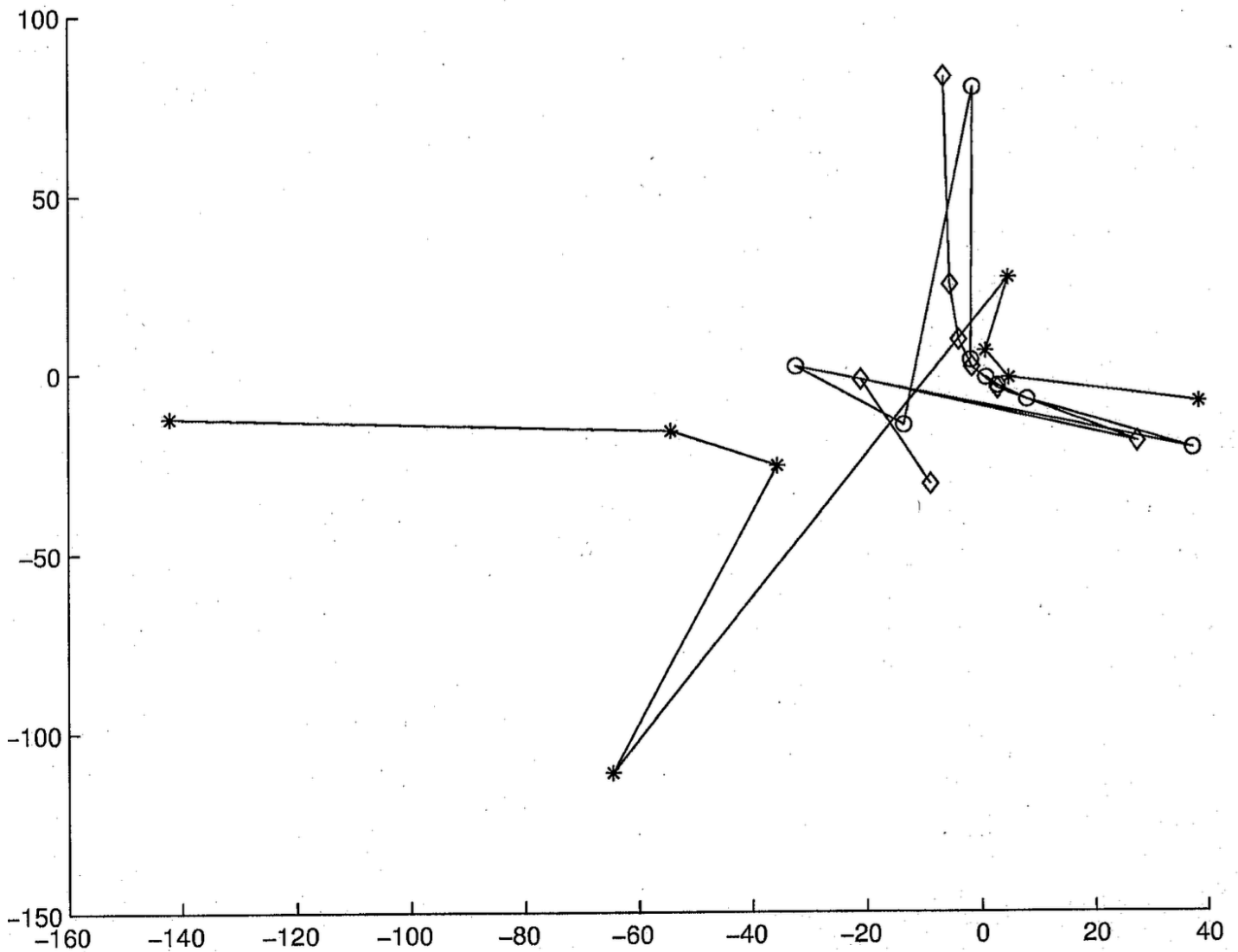


Fig. 19. For the last eight images in Fig. 18, the vanishing points for the cube's three sets of parallel lines are plotted as their image coordinates from Fig. 18. Their locations and how they vary with the cube's small rotations is "simple" mathematically (intersection of image lines), and the image determines the vanishing points and vice versa. But the point is that the relationship of vanishing points to rotations is not easy to visualize.

8. Appendix 3: Foreshortening (Section 2.1)

A3.1 Motivation and Goals

Here we shall see why this quote from a figure-drawing text is misleading.

When an object is tipped toward you or away from you it is said to be "foreshortened." It seems to diminish in size and change in shape as it recedes. In the case of the cylinder, notice that the sides become shorter and the ellipses more open as the foreshortening is increased.

These are the only definitions provided, indeed the only words about, perspective and foreshortening in that book, in a section entitled 'Foreshortening'.

I want to separate foreshortening from perspective and say a bit about its uses. Notably,

- * Foreshortening means that a slanted-away surface in the image appears smaller from our point of view than it does when viewed “head on” (when it is parallel to the image plane). But that is not due, usually, to how we image the scene (say, in perspective). The shrinkage with respect to a point of view occurs in the 3-D scene itself before any imaging takes place.
- * We can have foreshortening without perspective,
- * but we can't have perspective without foreshortening.
- * In usual (finite focal length) point projection, all surfaces are in perspective almost everywhere — Because in physics and 3-D geometry, all surfaces are foreshortened almost everywhere.
- * 3-D object shape determines the angle a surface makes with the gaze direction and with the direction of the light. The first angle determines foreshortening, the second is important in shading. Object shape and nature thus create shaded objects in space, which may then be imaged with or without perspective. (as in “modelled drawing”).
- * Both "modelled drawing" (with shading) and foreshortened patterns on a surface (“textured drawing”) allow reconstruction of the three-dimensional shape of the subject.
- * When both effects are operating (the usual case) there are even more clues to perceiving (or reconstructing) the solid shape of the subject, but we don't touch on that here.

A3.2 Foreshortening Without Perspective and Vice-versa

Fig. 19 shows orthographic projections of a cylinder and a sphere (hard to tell on sphere, but long axes of all polka-dots should be the same length). Recall that orthographic (parallel) projection is just point projection with an infinite focal length telephoto lens (the projection lines run perpendicular to the image plane, so objects do not shrink with distance.) We see this projection in blueprints and architectural drawings.

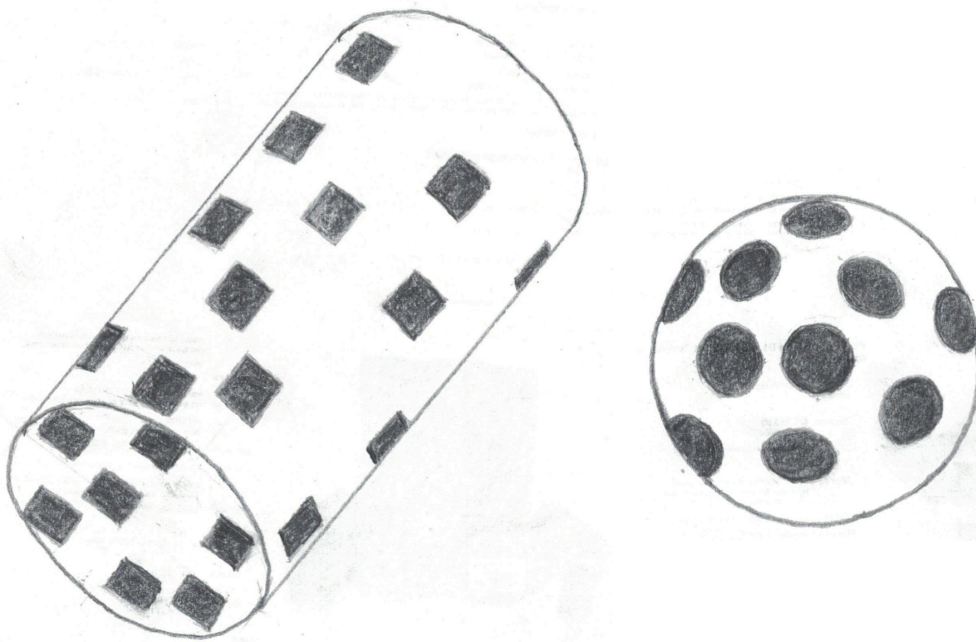


Fig 19: Non-perspective views of cylinder and sphere, with patterned texture. In cylinder, curvature of texture element sides is suppressed to show just foreshortening effect.

We see that the cylinder's sides must appear shrunken compared to a side-on view but that it's not shrinking with distance. If we imagine the image of a playing card, it is at its largest when we view it "head on" (gaze direction is perpendicular to its face or back). As it slants more and more, its image is compressed into a smaller size in the direction of its slant, until ultimately we see it edge-on as a straight line, leaving perspective totally aside. So we can have foreshortening without perspective.

In Fig 19, the varying foreshortening of texture elements, without shrinking in the distance, tells us about the shape of the objects.

The *vice-versa* case does not operate: That is, if a surface shows perspective shrinking, it must be slanted away from us, meaning it is also foreshortened. If it shows both effects, which is the usual case, and which is what confused the author of our opening quote. So perspective implies foreshortening but not the other way.

Most real-world objects have a set or range of surface orientations to the image plane. E.g. a sphere has them all. Surfaces facing away are, in a solid object, obscured by matter in front of them. Most real objects have at least one point (or if you like, one small area) parallel to the image plane that is unaffected by either perspective or foreshortening: The sphere in Fig 19 has one such point at its center. The ideal cylinder in Fig 19 has no such point, but a real manufactured cylinder's "sharp" edge may actually be rounded at some microscopic scale so as to create one.

A3.3. Foreshortening and Modelled Drawings

Modelled drawing simply takes advantage of (copying or inventing) *shading* ('shadowing' of varying intensity), to give the impression of depth.

It is the orientation of object surfaces that causes foreshortening. The physics of light, the angle it falls on the surface, the angle of the surface to the gaze direction, and the physics of the surface (its reflection function) all work together to create shading that depends on surface orientation. This shading is viewpoint-dependent but happens before any imaging is done, so the same object, already shaded, if you will, by nature, can be imaged (drawn) with or without perspective.

Humans are good at using the shading clue to attribute shapes to surfaces, and it is practically very important:

* The full moon looks flat since it doesn't have a surface that acts, say, something like a ping-pong ball's, so the full moon does not betray its shape through its shading -- this led to some confusion in the dim past.

* The make-up industry accounts for rather a lot of consumer spending, and makeup is often used to change perceived shape: to narrow the nose, emphasize (or create) cheekbones, reshape or create eyebrows...

* In modelled drawing, a well-respected text emphasizes the artist's freedom (if not duty) to discard or change shading that is not telling the shape story correctly. Cast shadows (shadows cast onto an already

shaded object), for instance, should often be ignored and removed, since they cover up and confuse the shading of the subject surfaces of interest. Multiple light sources cause confusion for similar reasons and should often be ignored, with “the right” modelling substituted. If the lighting changes (over time, say), the artist must ignore the current scene and make up shading consistent with the earlier time the drawing started. And so forth.

To see how foreshortening can give the same shape clues as shading, see Fig. A1 again, or imagine a surface with texture (herringbone, hairy skin, stripes, plaids, polka-dots, bricks...). It's easy to see that a textured drawing can be at least as informative about shape as a modelled drawing, and furthermore is not influenced, as is shading, by changes in lighting, multiple light sources, etc.

A3.4. Computer Graphics and Computer Vision

All the work described in this section involves geometry and physics, but not projective imaging: perspective is not involved in any of it.

Computer graphics has become increasingly sophisticated at modelled and texture “drawing”. Graphics starts with a geometric 3-D model of the object and a description of its surface properties (e.g. color, and how it reflects light hitting it from various directions: is it like a pingpong ball, or a mirrored ball in the garden, or one with semigloss paint?) The reflectance function interacts with the object’s “color.” For instance, under white light, an apple or a plastic object has white highlights while a gold ring has gold highlights. It gets complicated.

Starting from simple plastic-looking models of how surfaces reflect light, graphics techniques have gradually extended to copying the look of metals, cloth, rough textures, etc. Today (2016) hair probably is a solved problem, but human skin is still a challenge. Universities used to do graphics research, but now...

One of the goals of computer vision in the 1970's and 80's was "inverse graphics:" the computer's input is a digitized image, and its output should be a three-dimensional model. So it needed to know the same physics the graphics researchers did, and the same reflectance models, etc.

Now "inverse" problems are typically much harder than "forward" ones: Breaking a lightbulb is forward: putting it back together is inverse. Knowing where your fingertip it is from the angles and lengths relating shoulder, elbow, wrist, and finger joints is forward. Figuring out the angles from the tip location is inverse: for instance, there can be several or no solutions. Physics and optics solve the forward problem of

creating an image. How to solve the inverse problem of understanding an image?

Physics (and graphics) create "shading from shape". Computer vision wanted the inverse: "shape from shading". It made assumptions about object smoothness; unvarying lighting; a known, usually point, light source shape; known reflectance function (an exact mathematical model of flat or semi-gloss paint, for example). The resulting procedures (algorithms) could indeed take an input image and create an output "depth map" or 3-D shape descriptions, which were often moderately good approximations to reality.

A3.5 The Moral: Imaging (e.g. Perspective) is Independent of Nature (e.g. Shading)

Fig. 20 (below) is used without permission from, but with thanks to, M.C. Escher (*The Graphic Work*, English edition 2016, Taschen Books Köln). This little scene has shading resulting mostly from the angles of walls to the light, and cast shadows. Once the artist MCE arrives and picks a viewpoint, geometry dictates how the angled walls, windows, etc. are foreshortened. So far, all but viewpoint choice is done by Nature, Physics and Geometry.

Now MCE chose a point projection model, in fact two-point perspective. That modified the shapes of the foreshortened surfaces. He might have chosen a different projection — but wait, he did! Next he substituted the result of applying a distorting lens to part of the two-point perspective image.

The moral here is simply that foreshortening and shading are fixed by Nature in a given scene, and then (after that) the drawing is up to the artist. He can ignore the given shape or shading, which is one rather radical approach; or modify the shading (Section A3.3) to improve the representation of shape, or put the shaded and foreshortened scene into some perspective scheme, or use a nameless perspective and local fish-eye re-projection as in *Balcony* (Fig. 20). How he depicts nature is up to him, but foreshortened shape and the shading resulting from it are not artificial constructs: they are given by Nature for the artist to ignore, modify, or accept.

Below: Fig. 20. *Balcony*, by M.C. Escher.

