WHY PERSPECTIVE IS DIFFICULT: HOW TWO ALGORITHMS FAIL*

John R. Kender

Computer Science Department
Columbia University
New York, NY 10027

ABSTRACT

Attempting to derive image algorithms solely under orthographic projection is deceptively easy. However, orthographic algorithms often fail when applied to the perspective case. More critically, since many things simplify under orthography, such algorithms often give no evidence for their proper extension. This paper gives two such examples, showing both the problems that arise for them under perspective, and the suprising extensions that they require.

1. THE EDGE CURVATURE-GRADIENT SPACE RELATION

Given two planar surfaces meeting at a common edge, there is a particularly simple but highly useful relationship between their gradient space representation and the concavity or convexity (with respect to the observer) of their joining edge [3]. It forms the basis of much qualitative and quantitative shape recovery work [1]. It is rederived here, using some of the simplifying methods described in [2].

As it stands, this edge curvature rule fails under perspective. Newly rederiving it under a perspective imaging geometry is useful for two reasons. First, it analytically extends some of the existing theory concerning the dual-space nature of the gradient space to case of perspective geometry. Secondly, it indicates some of the difficulties that can occur in extending other orthographic relationships.

1.1. Under Orthography: The Rule Rederived

Since edge curvature is a relative measurement of the scene, both the rotational coupling of the gradient space and the translation invariance properties of the back-projection under orthography can be employed. All such problems can then be solved in the single standard position illustrated in Figure 1-1.

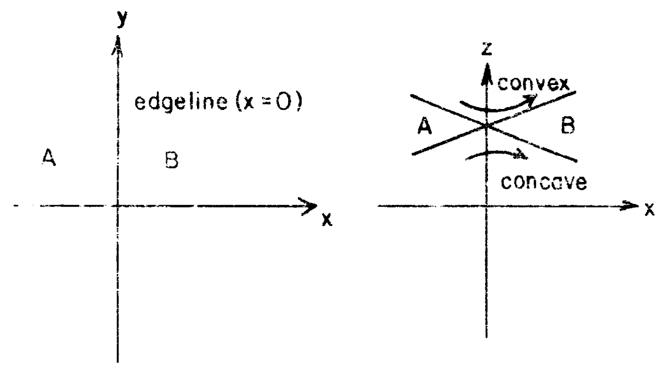

Figure 1-1: The standard configuration for analyzing curvature of edges under orthography.

The two planes are of the form $z=p_1x+q_1y+c_1$ and $z=p_2x+q_2y+c_2$. Since the edge they meet on is in the plane $x=0$, their simultaneous solution in that plane requires that $q_1=q_2$ and $c_1=c_2$.

The equality of the q values implies that the gradients of the two planes must lie on a horizontal line in the gradient space. Given the rotational coupling of the space, this is the equivalent of stating that the line joining the gradients is perpendicular to the image line. This confirms analytically the somewhat more elegant derivation of the same result by Mackworth, based on the dual nature of the gradient space.

If the planes meet to form a convex edge, then that edge will have a convex profile when cut by any plane not parallel to the edge itself. Similarly, concave edges generate only concave profiles. It follows that the converse is also true: a given profile also uniquely determines the curvature of the edge itself.

Consider then the profile of the two above planes determined by a slice along the plane $y=0$. The profile, in Figure 1-1, shows the cross sections $z=p_1x+c$ and $z=p_2x+c$, where c is the common c value. There are four possible physical arrangements of two half-planes in three-space. However, only two arrangements give rise to an image having two visible planes, one on either side of the edge line. These are also shown in Figure 1-1.

---

*This research was sponsored in part by the Defense Advanced Research Projects Agency under contract N00039-82-C-0427.

For the upper arrangement to be convex requires that the right half-line have a larger slope that the left; that is, $p_1>p_2$. Further, by the imaging geometry, the value p_1 must have been derived from the equation describing the right hand surface. It follows that a convex edge generates the constraint that the p value of the right-hand surface must exceed that of the left-hand surface: they are in the same relative left-right order in the gradient space as the surfaces are in the image. This ordering relationship is independent of orientation or translation. The analogous relation for concavity also holds, derived from the bottom arrangement of half-planes. Again, $p_1>p_2$, but the imaging geometry requires that p_1 is derived from the left-hand plane. Under this arrangement, the relative ordering of gradient values is reversed. Thus, the rule of Mackworth has been analytically rederived.

1.2. Under Perspective: The Rule Fails

The edge curvature relations do not hold under perspective, as the illustration in Figure 1-2 shows. In it, a bird's-eye view of the imaging situation is given, along with the perspectively distorted image that results.

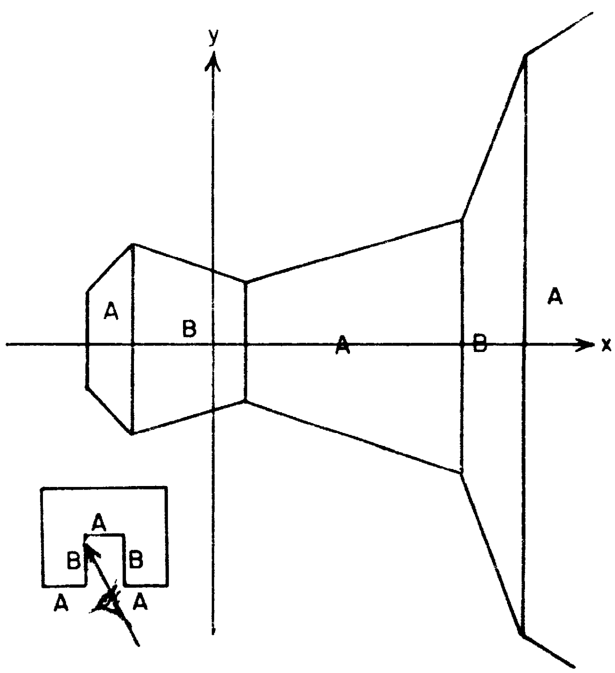

Figure 1-2: The failure of the edge curvature rule under perspective.

Suppose the edge curvature rule applied directly. Since all the surfaces labelled with like letters are parallel, they map into the same gradient point. (There is enough information in this image so that, with minimal assumptions concerning parallelism, the exact surface orientation of the surfaces can be analytically derived; see [2].) The gradient space thus consists of only two points, A and B. Since every edge is parallel to the y axis, the two points must lie on a horizontal line in the gradient space. However, it is impossible to assign a value to either point A or B and still uphold the edge curvature relation. This is because the image contains all four possible combinations of either concave or convex edge, and either left-right or right-left image ordering of A and B; any assignment would violate two of them. (Under orthography, the rightmost plane B would not be visible, being obscured by the rightmost A; the edge curvature rule would then apply.)

1.3. Under Perspective: The Rule Repaired

Although the rule fails, the cause is not immediately clear. It does, however, seem to be related to the fact that some of the surfaces that are invisible under orthography can be seen under perspective, due to non-parallelism of the imaging rays.

The edge curvature rule is again derived here, this time in the context of perspective, following the above derivation under orthography. Note that under perspective, only the rotational coupling of the gradient space can be employed, however; the result will necessarily be dependent on the position of the edge line in the image. The standard position is illustrated in Figure 1-3.

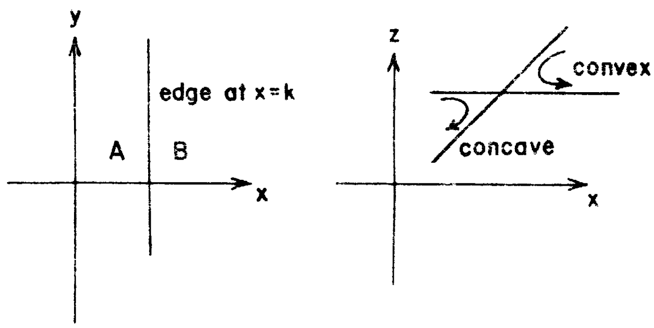

Figure 1-3: The standard configuration for analyzing curvature of edges under perspective.

The two planes are again of the form $z=p_1x+q_1y+c_1$, and $z=p_2x+q_2y+c_2$. Since they meet on the picture line $x=k$, they must meet within the image plane $x=kz$. The equation of their intersection is straightforward but complex. It can be written in the two-point form of a gradient space line as $(q_1-0)/(p_1-1/k) = (0-q_2)/(1/k-p_2)$. This implies that the gradients of the two planes must lie on the same line through $(1/k,0)$ in the gradient space; that is, through the dual (the gradient) of the plane $z=(1/k)x$. (As perspective degenerates to orthography, k approaches zero, and the gradients lie on a line parallel to the p axis: the result under orthography.) Again, this confirms the result of Mackworth, derived using the dual nature of the gradient space.

Further, the plane $z=(1/k)x$ is the one member of the parallel family $z=(1/k)x+\emptyset y+c$ that passes through the origin: it forms their common vanishing line by its intersection with the picture plane. This plane through the origin is the "interpretation plane" of Mackworth [3]. Thus, the dual of the interpretation plane of the edge under perspective can be obtained by considering the edge to be a vanishing line. If the edge line is written as $G.P=1$, its dual is $G/||G||^2$.

Now consider the profile of the two surface planes determined by a slice along the plane $y=\emptyset$. Figure 1-3 shows the cross sections. As before, the curvature of the profile determines the curvature of the edge, and only two arrangements give rise to an image having two visible planes. Note, however, the critical difference under perspective: edge curvature must be measured with respect to the <u>line of sight</u> to the edge, as the imaging rays are no longer parallel to the y axis.

Explicit analysis can be avoided by means of the following geometric argument. The gradient space is panned about the y axis so that the edge lies on the positive z axis, the gradient relations are established in the new coordinate system, and then the gradient space (with the relations) is panned back. The angle of pan is given by $\tan\Theta = k$. Having the edge on the z axis is the orthographic case.

Suppose the edge is convex. In the panned gradient space, the gradient of the right-hand plane must be greater than the left-hand plane. However, when the gradient space is panned back, the relation necessarily has a different form, given that one or both gradient points may have "wrapped around through infinity". The gradient of those planes with infinite p value in the panned coordinate system also wrap around, to the point $(1/k,\emptyset)$.

Therefore, the edge curvature rule still holds if the gradient space is considered to be "cut" into two half-spaces at $(1/k,\emptyset)$, and its two infinite ends "sutured" together. This alters the concept of order along the line through $(1/k,\emptyset)$. Due to the cut, B "is to the left of" A is considered to be true under any cyclic permutation of the ordering: B, A, cut. Thus, A cut B is valid, but B cut A is not. (Under orthography no suturing is needed.)

The ambiguity of labeling the gradient points in the example given in Figure 1-2 is now resolvable. For example, consider the rightmost convex edge, located L units from the image origin. It still obeys the edge curvature rule, given that the gradient space is cut at L's interpretation plane's dual, $(1/L,\emptyset)$. Under the suturing, B (which has a positive gradient greater than $1/L$) "is to the left of" A (which has a negative gradient), since the value of the cut lies between them. This is the case: A cut B.

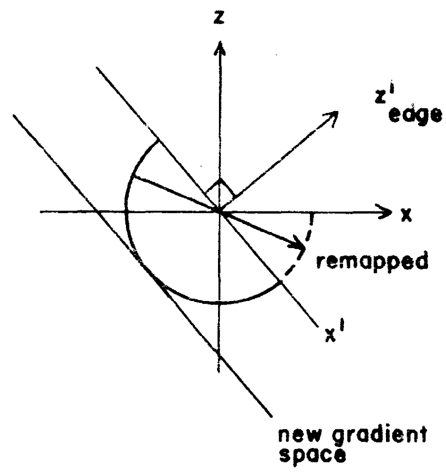

Figure 1-4: The repair of the edge curvature rule under perspective.

These relations can be more easily illustrated on the Gaussian sphere. As seen in cross-section at the plane $y=\emptyset$ in Figure 1-4, the panned gradient space corresponds to a partial revolution of the Gaussian sphere about the y axis. Planes of infinite p with respect to the line of sight are located in the standard coordinate system at $(1/k,\emptyset)$. Wrap-around is more easily seen: standard gradient points of higher p value than $1/k$ are really the remapping of slopes "past" negative infinity. These values are actually the projection of part of the "front" surface of the sphere.

In any case, the form of the edge curvature rule under perspective is non-trivial, and would complicate the perspective extension of existing orthographic algorithms employing it.

2. <u>ONE OTHER RELATION DIFFICULT TO EXTEND</u>

A second example further demonstrates that it is often difficult to predict the perspective form of a relation from its orthographic counterpart. Additionally, it indicates that what has been simplified under orthography may yet have a simple form under perspective.

Consider the intersection of several planes at a common vertex. Under orthography, the dual of the interpretation planes of each edge is important: this dual constrains, by the left-right rule, the choice of gradients for the pair of planes that flank it. However, each dual is infinite, since under orthography the interpretation plane is parallel to the line of sight. There is therefore no apparent relationship among the duals.

However, under perspective, the duals of the interpretation planes are usually finite; they can be determined as if the edges were the vanishing lines of the interpretation planes. (The exception occurs when the edge passes through the origin, implying an infinite gradient.) Further, as Figure 2-1 shows, the duals all lie on a common line in the gradient space. If the vertex is at picture point P, the line is simply G.P=1. Thus, this line is perpendicular to the vertex-picture origin direction, and at a distance of 1/||P|| from the gradient space origin.

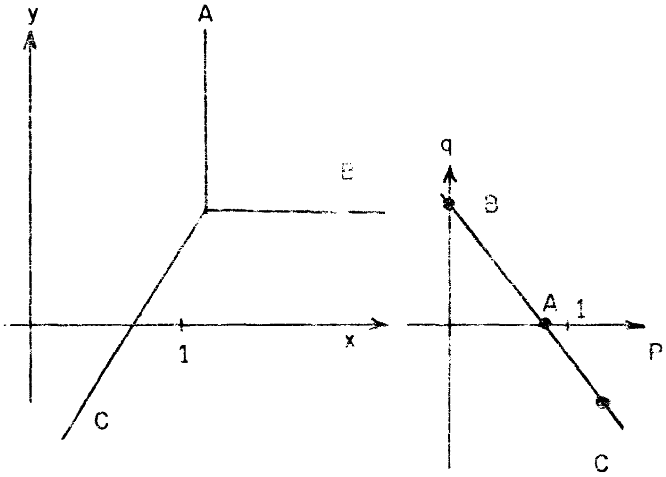

Figure 2-1: Interpretation plane duals, under perspective, are finite and well-behaved.

The explanation is not difficult. All the interpretation planes must be parallel to the line of sight to the vertex; this is the vector T=(x,y,1). Planes parallel to a given vector are those planes perpendicular to the plane whose normal is the vector. Thus, interpretation planes are of the form N.T=0, or equivalently, G.P=1.

This line of duals can simplify the problem of assigning the proper configurations of gradients for the planes AB, AC and BC. However, like many things under perspective, it has disturbingly non-linear behavior under translation.

3. WHY PERSPECTIVE LINE DRAWINGS ARE DIFFICULT

Thus, there are at least two problems associated with extending known orthographic results in line drawing analysis to perspective.

The first is that the gradients of planes meeting at a common edge no longer obey a simple relation. Related planes lie on a line through the dual point of their interpretation plane (the point (1/k,0), appropriately rotated). This is more cumbersome than having them lie on a gradient space line perpendicular to the image line. Further, as the edge moves in the image, this common point moves, too, in a non-linear way; under orthography, there are no changes.

Secondly, concavity and convexity have no simple gradient space analogues, either. The gradient space must be viewed with respect to the "cuts" created by line-of-sight effects. Again, as an edge moves, the cut moves, and the relationships are also affected; not so in orthography.

Although the line-of-duals relationship discussed above can be an aid to line-drawing analysis--an aid unnecessary under orthography--it still appears that any extension of existing line-based techniques to the perspective case will be difficult and inelegant. Critical here is the strong translation dependence of perspective. In part, this gives further reason to encourage a trend to more surface-based approaches in image understanding work.

REFERENCES

[1] Kanade, T.
A Theory of Origami World.
Artificial Intelligence 13(3):279-311, May, 1980.

[2] Kender, J. R.
Shape from Texture.
PhD thesis, Computer Science Department, Carnegie-Mellon University, 1980.

[3] Mackworth, A. K.
Interpreting Pictures of Polyhedral Scenes.
Artificial Intelligence 4(2):121-137, Summer, 1973.