MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 1555 October 1995

The Three-Dimensional Interpretation of a Class of Simple Line-Drawings

Thomas Marill

Abstract

We provide a theory of the three-dimensional interpretation of a class of line-drawings called p-images, which are interpreted by the human vision system as parallelepipeds ("boxes"). Despite their simplicity, p-images raise many interesting vision questions:

• Why are p-images seen as three-dimensional objects? Why not just as flat images?

• What are the dimensions and pose of the perceived objects?

• Why are some p-images interpreted as rectangular boxes, while others are seen as skewed, even though there is no obvious distinction between the images?

• When p-images are rotated in three dimensions, why are the image-sequences perceived as distorting objects—even though structure-from-motion would predict that rigid objects would be seen?

• Why are some three-dimensional parallelepipeds seen as radically different when viewed from different viewpoints?

We show that these and related questions can be answered with the help of a single mathematical result and an associated perceptual principle.

An interesting special case arises when there are right angles in the p-image. This case represents a singularity in the equations and is mystifying from the vision point of view. It would seem that (at least in this case) the vision system does not follow the ordinary rules of geometry but operates in accordance with other (and as yet unknown) principles.

1 Introduction.

Line-drawing analysis has received a substantial amount of attention in the last thirty years. In an informative capsule review of the field, Horn [1986, p. 360-362] points out that "the analysis of line drawings was at one point the focus of vision work in the artificial intelligence community."

This interest may be due to the fact that analyses of line-drawings start with symbolic representations rather than images. These analyses therefore bypass the field of "early vision" and concentrate instead on the "later" (and perhaps more fundamental) aspects of the vision process, in particular as regards three-dimensionality.

How do we define the problem to be solved? The definition has been approached in two ways:

(a) The recovery approach. In much of the literature on line-drawing analysis (indeed in much of the literature on computer vision), the problem is taken to be the problem of recovering the object or scene that generated the image.

(b) The psychological approach. Alternatively, the problem can be defined as the problem of finding an interpretation of the image that matches the interpretation generated by the human vision system.

In the present work, we use the second definition. Various reasons for rejecting the recovery approach are given in Sections 8 and 9.

Either way we look at it, and despite the many excellent contributions that have been made to the field, the problem of line-drawing analysis has not been solved. The startling fact is that even today there is not in existence a single program that can accept a wide range of line-drawings and produce satisfactory output under either definition of the problem. By some reckoning, we may not even be close. Even the very simple images considered in the present work have not hitherto been handled satisfactorily.

2 Parallelogram meshes and p-images.

A *parallelogram mesh* is a planar configuration consisting of one or more parallelograms; each parallelogram in the mesh shares one or more sides with other parallelograms. In Figure 1 we see examples of randomly-constructed parallelogram meshes.

A *p-image* is a specific type of parallelogram mesh, consisting of six parallelograms; each parallelogram shares each of its sides with one other parallelogram. In Figure 2 we see examples of randomly-constructed p-images. We note that all of these p-images have the same number of angles, lines, and points. Only the lengths of the lines and the measurements of the angles differ.

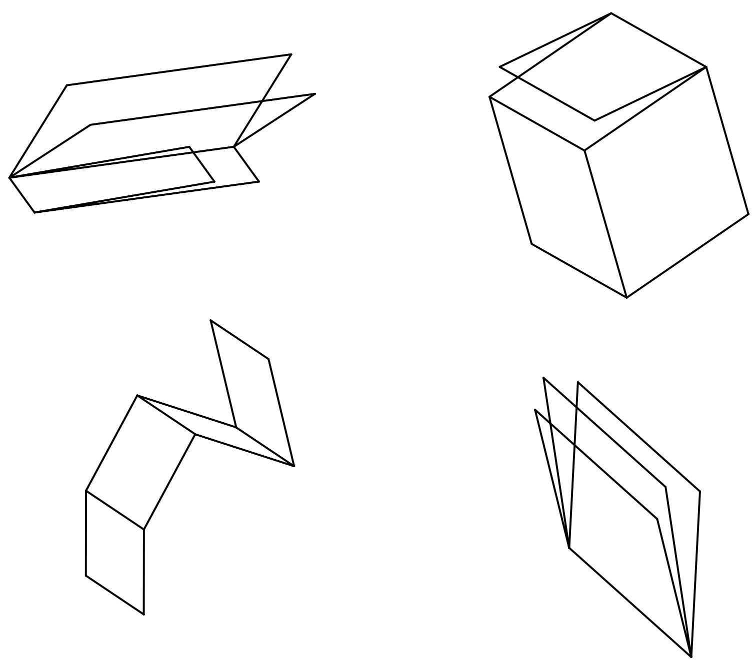

Figure 1: Random parallelogram meshes.

A p-image is determined by any one of its triple vertices (as defined in Section 3). We can think of the three lines of such a triple vertex as forming the "basis vectors" of the p-image (Figure 3); given these vectors, we can construct the p-image.

P-images are interpreted by the human vision system as three-dimensional

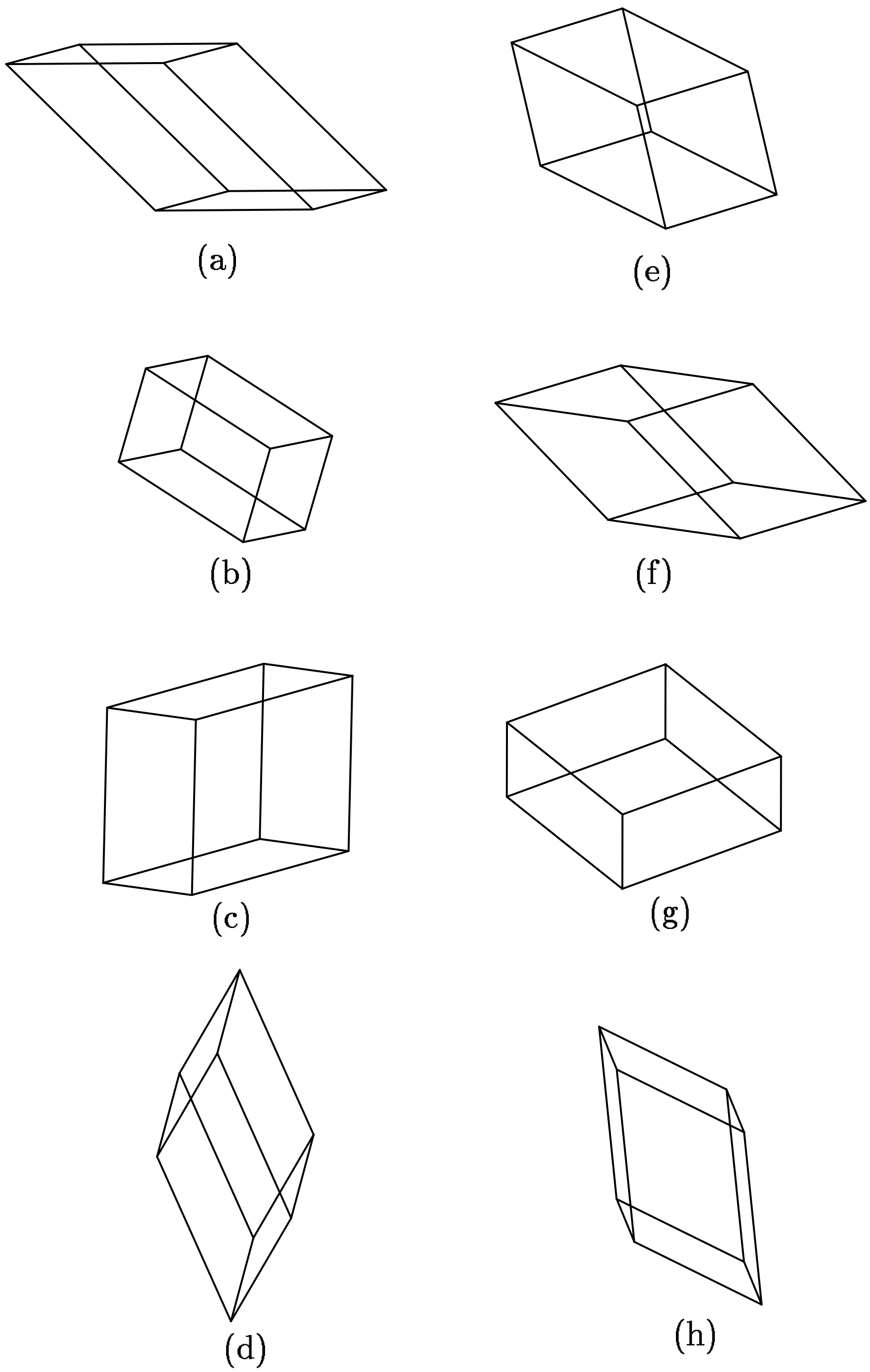

Figure 2: Random p-images.

transparent boxes (parallelepipeds). If we remove one of the interior triple vertices from a p-image (Figure 4), we obtain a *reduced* p-image, which is interpreted visually as an opaque box. With the exception of this transparent/opaque difference, all the results we shall discuss apply equally to reduced or non-reduced p-images.

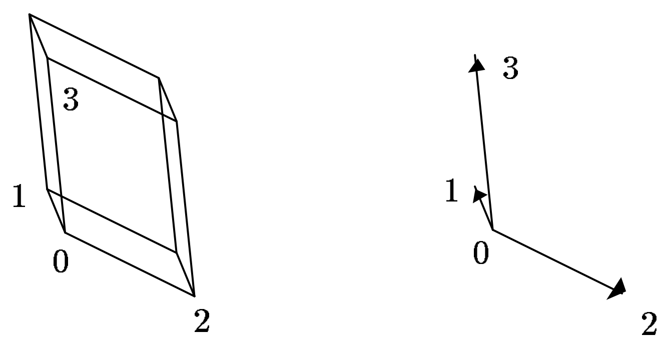

Figure 3: Basis vectors for p-image.

3 Conforming and non-conforming p-images.

The p-images of Figure 2 can be divided into two classes. Images (b), (c), (e), and (g) are all interpreted by the vision system as rectangular boxes (parallelepipeds having rectangles as sides). Images (a), (d), (f), and (h) are interpreted by the vision system as skewed boxes (parallelepipeds having non-rectangular parallelograms as sides).

This separation into two classes is a puzzling phenomenon, since there is no immediately obvious difference between the images in the two classes. One might perhaps guess that images interpreted as skewed boxes have a larger number of acute angles than the images interpreted as rectangular boxes. But, in fact, all p-images have the same number of acute angles[1].

[1]P-images with angles of ninety or zero degrees represent special cases. The case of ninety degrees is considered in Section 9. The case of zero degrees is not considered here.

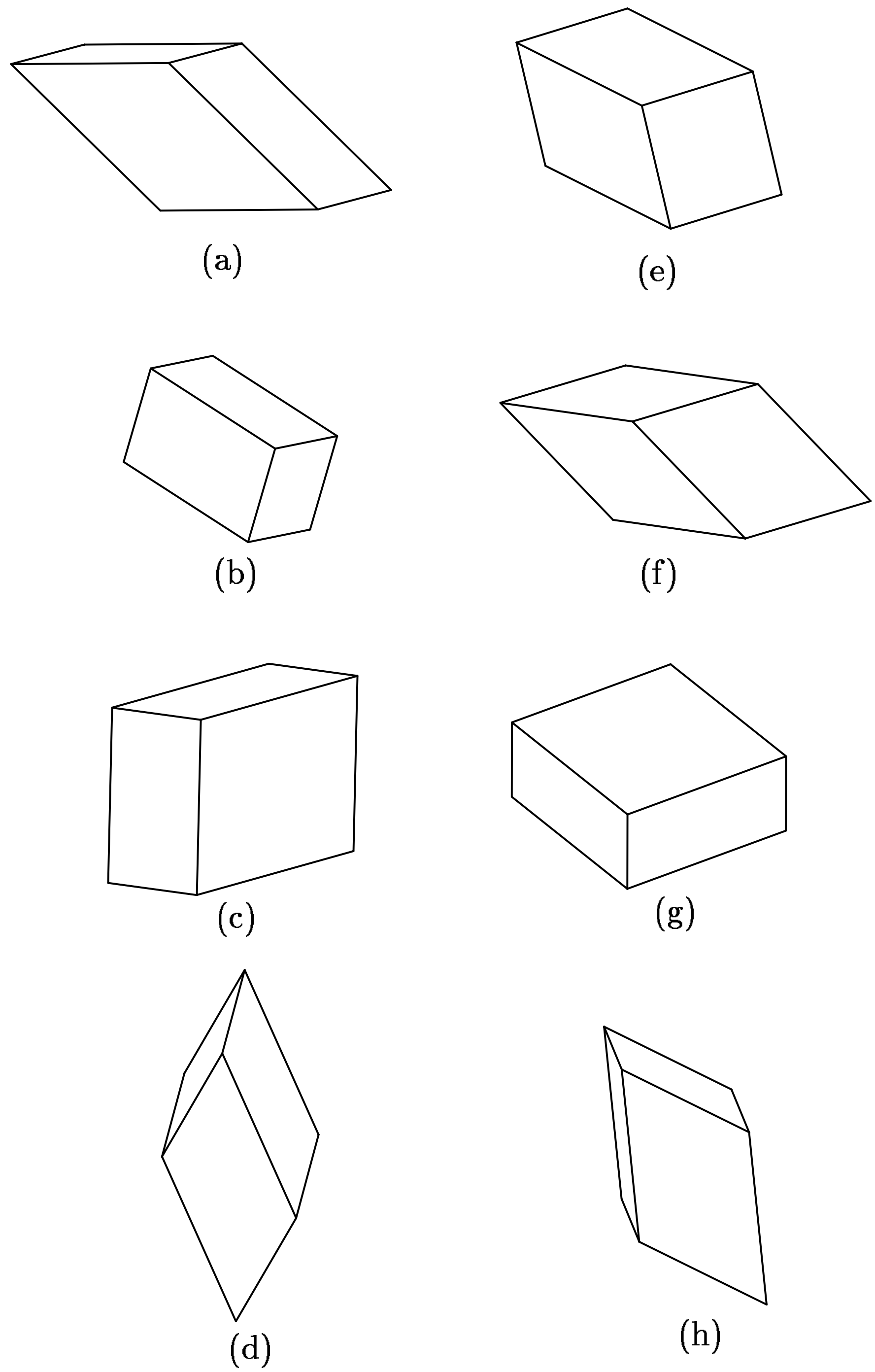

Figure 4: Reduced p-images corresponding to Figure 2.

Or one might guess that the images interpreted as rectangular boxes have right angles. In fact, none has a right angle, or an angle close to a right angle. Image (g), for example, has no angle that comes within 20 degrees of a right angle.

Or again it might be guessed that the images that are interpreted as rectangular boxes receive this interpretation because they in fact *are* the projections of rectangular boxes and the vision system somehow correctly detects this fact (similarly for the skewed boxes). But a moment's reflection will tell us that this cannot be the case. The images seen as rectangular boxes can in fact be the projection of boxes that don't even have parallel edges. And even if we restrict ourselves to boxes with parallel edges, we will see in Section 8 that three-dimensional skewed boxes can project to p-images that are perceived as rectangular.

What then accounts for the difference between the two classes of images?

A considerable amount of light is shed on this question by the following result, proved in Appendix A.

Definition. A *triple vertex* is a two-dimensional or three-dimensional configuration in which three line-segments coterminate at a point to form three angles.[2] A *right-angled triple vertex* is a three-dimensional triple vertex having three right angles.

Theorem. Every planar triple vertex V is the (orthographic) image of some right-angled triple vertex, unless V contains a right angle or an odd number of acute angles.

It can now be observed that images (b), (c), (e), and (g) of Figures 2 and 4 conform to the conditions of the theorem. That is to say, none of the triple vertices in these images contains a right angle or an odd number of acute angles. We refer to such p-images as *conforming* images. Furthermore, we observe that images (a) (d) (f) and (h) of Figures 2 and 4 fail to conform to the conditions of the theorem (all of the triple vertices of these images contain an odd number of acute angles). We refer to such p-images as *non-conforming.*

[2]Such a configuration is sometimes called a "trihedral angle". However, the word "trihedral" implies three surfaces, and is therefore inappropriate in the present context, in which we deal with lines rather than surfaces. The term "trihedral" also seems inappropriate for planar configurations.

4 A perceptual principle for parallelogram meshes

It is generally taken as a fundamental but unspoken axiom of vision that the visual interpretation of an image (in particular, a three-dimensional interpretation), must project to that image. Thus, if we define the set of objects that project to a given image as being the *extension* of that image, then this axiom states that the visual interpretation of an image must be in the extension of that image. It is for this reason that so much vision work deals with geometry.

This axiom is a necessary underpinning for the theory that vision "recovers" the object that generates the image. If the image is created by the object, then clearly the object is in the extension of the image. But if the interpretation is not in the extension, then the interpretation cannot be identical to that object, and the object cannot be recovered.

We will accept this axiom. (But see Section 9 for some serious second thoughts on this issue.)

This axiom and the theorem of Section 3 take us part of the way toward understanding the two classes of p-images. We can go further with the help of the following:

Perceptual principle: (a) Given a parallelogram mesh, the vision system will interpret all parallelograms as rectangles, if it is possible to do so. (b) If it is not possible, the system will interpret parallelograms as parallelograms.[3]

We note that in order for a parallelogram to be interpreted as a rectangle it must be rotated out of the image plane. Thus the perceptual principle is sufficient to explain why conforming p-images are seen as three-dimensional.

Of course, one can then ask the deeper question: why does the vision system act in accordance with such a principle? A possible answer has been suggested by Marill [1992]: the 3D rectangles are, in a certain mathematical sense, less complex than the 2D parallelograms and require fewer bits for their representation.

We can also now understand why conforming p-images are seen as having right angles. The perceptual principle tells us that the parallelograms in the p-image will be seen as rectangles if possible, and the theorem of Section 3 tells us that it

[3]It must be remembered that it is geometrically possible to interpret a parallelogram in the image as a quadrilateral in space having no parallel lines. In fact, the interpretation need not even be planar. Thus statement (b) is less tautological than it sounds.

is possible. Thus, a conforming p-image will be interpreted as a three-dimensional configuration of linked rectangles that projects to the image (in short a rectangular box).

With the mechanisms developed so far, we are not yet at the point of being able to predict the dimensions or the pose of the object that will be seen. Nor are we at the point of understanding what happens with non-conforming p-images. These issues are taken up in the next two sections.

5 The dimensions and pose of the perceived object; the conforming case

.

Consider the three basis vectors of a p-image (Figure 3). We saw in Section 4 that a conforming p-image will be interpreted as a rectangular box. Thus, in the 3D interpretation, the three angles formed by the basis vectors will be right angles.

Let us assume that the tails of the vectors are located at the origin and that we know the location in the image of the heads of the vectors.[4] We can then solve for the z-coordinates of the heads of the vectors in the 3D interpretation (see Appendix A), obtaining:

(1) $$z_1 = \pm\sqrt{-a_{12}a_{13}/a_{23}}$$

(2) $$z_2 = \pm\sqrt{-a_{12}a_{23}/a_{13}}$$

(3) $$z_3 = \pm\sqrt{-a_{13}a_{23}/a_{12}}$$

where the a_{ij} are known constants, defined in equations (13), (14), and (15) of Appendix A.

[4] Here, and throughout, we assume orthographic projection unless specifically stated otherwise.

Using these formulas, and given a conforming p-image, we can determine the dimensions and pose of a rectangular box that projects to that image. There are always two solutions, depending on whether one picks the three quantities as positive or negative. (The two solutions generate two objects that are mirror images of one another, reflected in the image plane.) Informal experiments show that the results obtained by this method are consistent with the interpretations of the human vision system.

6 Non-conforming p-images and the "compromise" heuristic.

The theory given above allows us to predict the interpretation, including dimensions and pose, of conforming p-images. But what about non-conforming ones?

Let us consider the three basis vectors. If we "anchor" the tails of the vectors at the origin, there are three degrees of freedom to be determined. In the conforming case we were able to get a solution by making all three angles into right angles.

We can interpret the perceptual principle of Section 4 as saying that the vision system "wants" to make right angles. But in the non-conforming case, the geometry does not allow all three angles to be right angles, since a triple vertex in a non-conforming p-image contains an odd number of acute angles, and the theorem of Section 3 tells us that such a triple vertex cannot be the image of right-angled tripled vertex. There is nothing, however, that prevents the vision system from making *two* right angles among the three.

But what about the third? We can proceed along the lines of the following "compromise" heuristic. With two right angles, there are still infinitely many possibilities for the three z-coordinates z1, z2 and z3. However, we can write z2 and z3 as functions of z1, and we can do this in several ways, depending on which of the angles are made into right angles. Let us pick two of these ways and find the value of z1 that minimizes the differences between these two ways. This yields a complete interpretation of the image. Such an interpretation is, in a sense, the best available compromise.

As we show in Appendix B, this approach yields the same equations (1), (2) and (3), as before, except that the sign under the radical is changed. Thus we can get complete interpretations of p-images in both the conforming and non-conforming cases by using a single set of equations, making sure we pick a sign under the radical that gives us real values. This fact simplifies and unifies the entire system.

But are the interpretations generated in this way the same interpretations that the human vision system generates? It is difficult to be absolutely sure. When looking at a conforming p-image, the visual interpretation is usually quite clear and precise. In the case of a non-conforming p-image, it is less clear; to describe the interpretation, one must form estimates of the lengths of lines or the magnitude of angles, something people are not good at. The best one can say is that results obtained by the above technique appear to be acceptable versions of the human interpretation.

Let us look at an example (Figure 5). The compromise heuristic interprets this image as a parallelepiped having four rectangular faces and two non-rectangular parallelogram faces. Faces 1-0-2-5 and 1-0-3-4 are both rectangles (these are shown separately in Figure 5(b)). In the interpretation, line 0-1 has length 3.6, line 0-2 has length 6.9, and line 0-3 has length 2.9. Angle 2-0-3 measures 66.7 degrees. All of this seems psychologically acceptable. Informal experiments with other examples yield similarly acceptable results.

(a) (b)

Figure 5: Interpretation of non-conforming p-image

7 Rotating p-images

A curious phenomenon occurs when we rotate a conforming p-image in three dimensions.[5] Suppose, for example, we take image (g) of Figure 2 and make a movie by rotating it in 3D around the y-axis. When we look at the movie, what we see is a *distorting* three-dimensional object that contracts and expands like an accordion; the object is not seen as rotating. An imperfect idea of what one sees in the movie can be got by looking at individual frames (Figure 6).

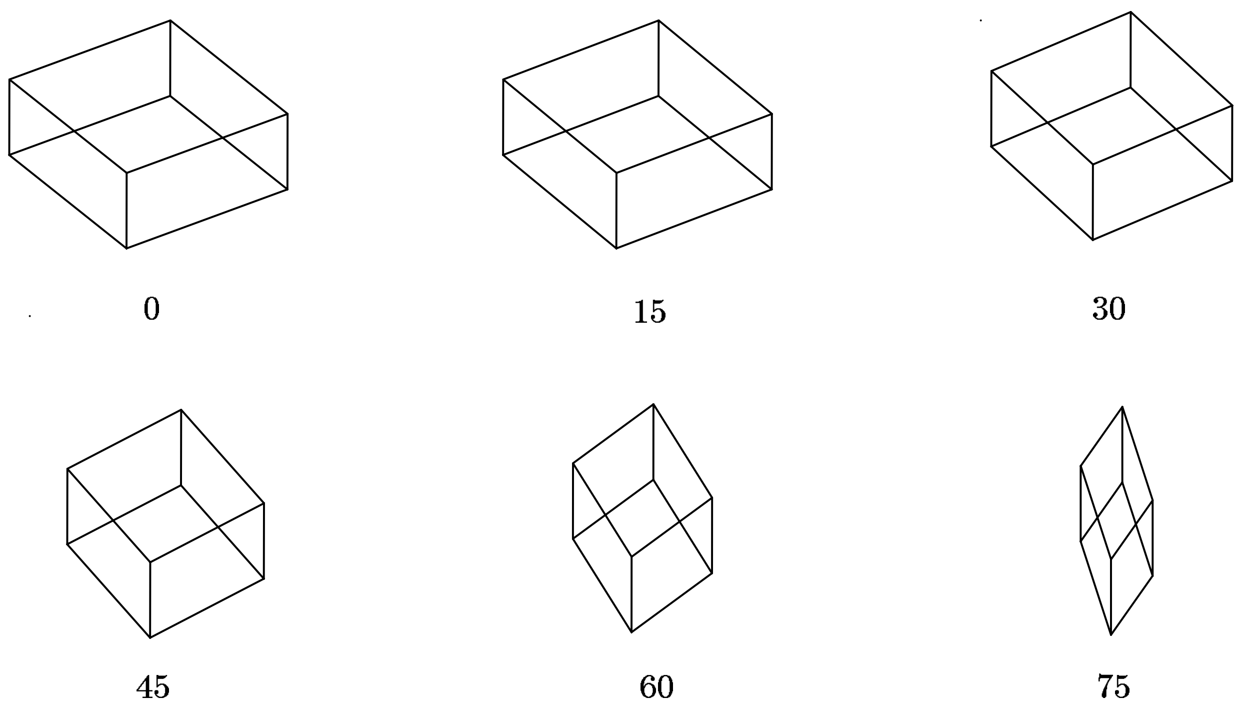

Figure 6: Rotated views of Figure 2(g). Angle in degrees.

We tend to believe that when we look at the movie of a rigid object in motion, we will see a rigid object in motion. Put another way, one believes that a time-varying image that is the projection of a rigid object in motion will be interpreted as a rigid object in motion. This belief underlies structure-from-motion theory [Ullman 1979].

In the present case, however, this belief does not hold true. The movie of our rotating p-image *is* the time-varying image of a rigid (albeit flat) object in motion, but it is not perceived as such. Instead, it is perceived as a deforming body that stays more or less in the same place.

[5] Note that we are rotating the *image*, not the box.

We are now in a better position to understand why. Take any frame in the sequence; it is a p-image, and our theory tells us exactly how it will be interpreted: it will be seen as a certain predictable three-dimensional box. The box will change shape in a predictable manner throughout the image sequence. The boxes will be rectangular up to a certain point in the sequence because the images are conforming up to that point; after that, the boxes will be skewed, because the images are non-conforming. The dimensions of the perceived box change in accordance with the predictions of the theory.

8 Paradoxical views of parallelepipeds.

Parallelepipeds project to p-images. Until now we have generated our p-images randomly. What would happen if we generated them by projection from three-dimensional parallelepipeds? Would the vision system, somehow, recover the objects that generated the images?

We tested this idea by using the three-dimensional parallelepiped specified in Appendix C (a skewed parallelepiped centered on the origin). We generated six views of this object by rotating the object around the y-axis, with rotation angles forty degrees apart.[6] The results are shown in Figure 7.

The six views are interpreted by the human vision system as six different objects. Some are rectangular boxes and some are skewed. Some are fat and some are thin. One of them looks like a cube, while others are greatly elongated. Thus the human vision system, for these images, does not come close to recovering the object that generated the images.

Our present theory predicts a different three-dimensional interpretation for each of these images. The predictions are in agreement with the interpretations of the human vision system.

[6] Note that we are here rotating the box, not the image.

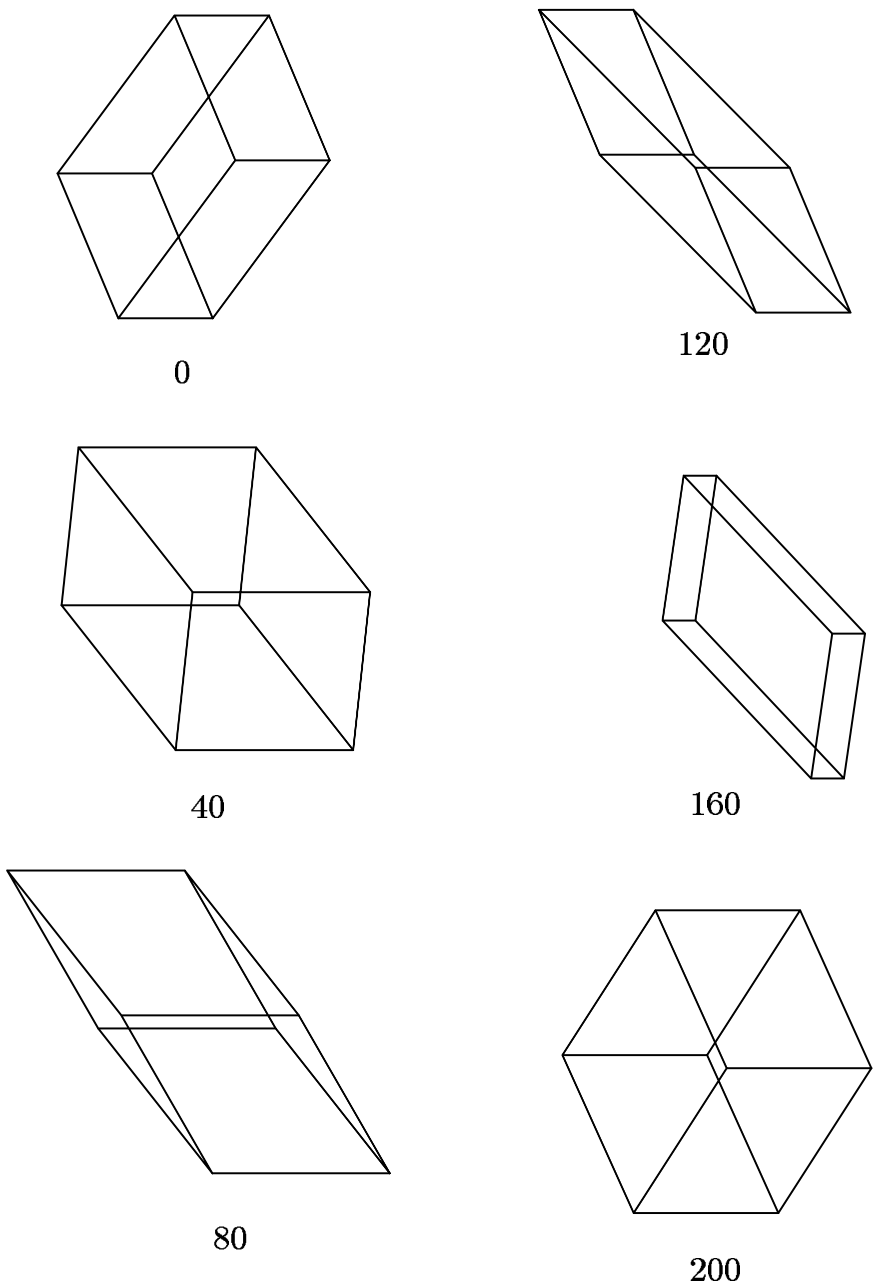

Figure 7: Different views of the same 3D parallelepiped. Rotation angle in degrees.

9 The special case of right angles: does vision follow the rules of geometry?

We have yet to consider the case in which there are right angles in the p-image. In this case, the value of one of the z-coordinates given by equations (1), (2), and (3) is undefined, since one of the denominators inside the radical is zero. Thus, the equations do not give us a solution.

What does the vision system actually do in this case? The answer is rather mystifying. Let us look at an example of a p-image with right angles (Figure 8 (a)). The image is interpreted visually as a cube.

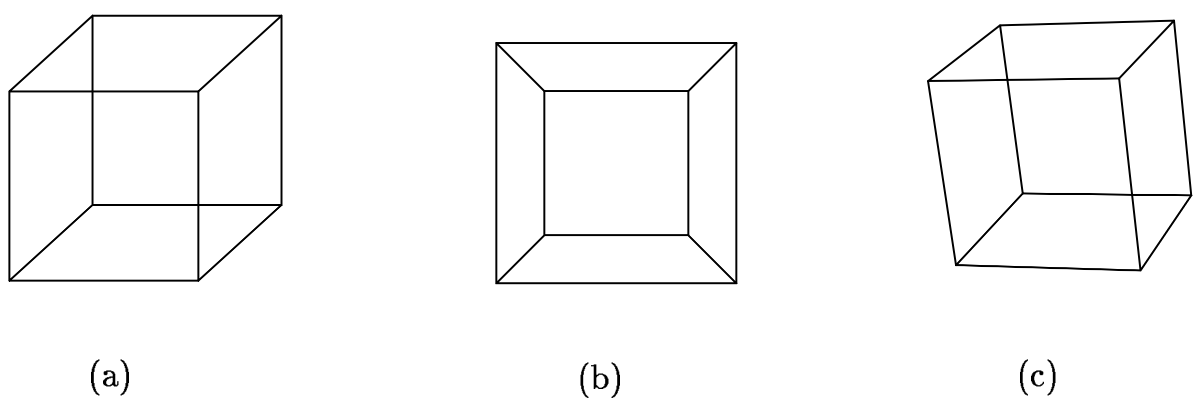

Figure 8: (a) A paradoxical p-image. There is no cube that projects to this image. (b) Perspective projection of frontal cube. (c) Perspective projection of tilted cube.

However, *there is no cube that projects to this image.* Under orthographic projection, if there were such a cube, there would be a planar triple vertex which contains a right angle and which is the image of a right-angled triple vertex in space, in contradiction of the theorem of Section 3. Under perspective projection, the image of a cube in general position has no parallel lines; see Figure 8(c). (At most, the perspective image of a cube has two sets of parallel lines; this case occurs when the front face of the cube is parallel to the image plane; see Figure 8(b).) In fact, however, Figure 8(a) has *three* sets of parallel lines.

In Section 4 we discussed what we called an unspoken axiom of vision that states that the visual interpretation of an image (in particular, a three-dimensional interpretation) must project to that image.

But the present example casts serious doubt on the validity of this axiom. It would appear that the visual interpretation of the image of Figure 8(a) is *not* in the extension of that image; i.e., what we see when we look at Figure 8(a) is not something that projects to Figure 8(a). It seems impossible to reconcile this observation with the idea that vision recovers the object that caused the image. The perceived object and the image are no longer related by the usual geometric rules that determine image formation, but by some other, as yet undetermined, set of rules.

Suppose we had a program that was to return a psychologically acceptable interpretation, given a line-drawing. What should the program return for Figure 8(a)? We know that there is no cube that projects to this image. We also know that there are infinitely many other 3D wire-frames that do. What shall the program pick?

The program could easily construct a cube-like wire-frame that projects orthographically to Figure 8(a), has square front and back faces, and has edges of equal length. But then the top and side faces would have angles of 20 and 60 degrees.

Alternatively, we could ask that the front and back faces be square and that all angles be 90 degrees. This can be approximated closely; but then the lengths of the edges will be greatly dissimilar. For example, we can make all angles within 0.02 degrees of right angles by making the edges of the front and back surfaces of length 3.6 and the other edges of length 6000.

This matter seems quite puzzling and worthy of further investigation.

10 Discussion of related work.

The concept of *skewed symmetry*, a property of a planar curve, was introduced by Kanade [1981]. Kanade proposed a principle according to which a skewed symmetry is interpreted as the projection of a real symmetry which is tilted out of the image plane; and he was able to show the relation between the skewed symmetry and the tilt of that plane. However, there are infinitely many tilted real symmetries that project to any given skewed symmetry; Kanade proposed that the correct interpretation is the one that minimizes the tilt.

Using this powerful principle, together with the theories of line-labeling (Clowes [1971] and Huffman [1970]) and of gradient space (Mackworth [1974]), Kanade was able to recover the 3D shape of objects from line-drawings in a number of cases, including the case of line-drawings of boxes similar to our reduced, conforming p-images. In the course of his analysis Kanade also proved, for the case of reduced p-images, that such images can be the projections of rectangular boxes if and only if the three angles in the interior triple vertex of the image are obtuse. This result, which applies to a certain kind of triple vertex called a "fork", is subsumed under our theorem of Section 3, which applies to any triple vertex.

Skewed symmetry does not help in the case of non-conforming p-images; in that case the vision system does not interpret skewed symmetries as tilted real symmetries (rectangles), as skewed-symmetry theory would require, but rather as non-rectangular parallelograms.

Kanade's approach has been criticized by Brady and Yuille [1983]. These authors state that Kanade's approach predicts that real symmetries will be interpreted as lying in the image plane, and they argue that this prediction is disproved by the case of an ellipse (which is a real symmetry, but is interpreted as a circle tilted out of the image plane). They propose a principle of their own for determining three-dimensional surface orientation from a planar contour: maximize the ratio of the area enclosed by the contour to the square of the perimeter of the contour.

However, Brady and Yuille are themselves open to a criticism somewhat similar to their criticism of Kanade: namely, that their compactness principle interprets a parallelogram as a slanted square, while the human vision system interprets a parallelogram as a slanted rectangle.

Brady and Yuille focus on the interpretation of single, closed planar contours. However, they claim that their principle also correctly interprets images such as the ones discussed here.[7] Friedberg [1986] disputes this claim. He points out that under Brady and Yuille's compactness principle each face of a perceived parallelepiped will be interpreted as a slanted square, but the orientation of three such squares at a vertex will not be consistent with the constraints derived from the shared edges because the faces of the object are in fact not square.

More recently Marill [1991] introduced the principle of "minimum standard deviation of angles" (MSDA) for the purpose of interpreting a wide class of line-drawings. (This idea, along with several other concepts for the interpretation of general line-drawings, had been suggested at an earlier date by Barrow and Tenenbaum [1981].) However, counterexamples to MSDA were found by Leclerc and Fischler [1992], who then proposed an enhancement to MSDA, whereby both the standard deviation of angles *and* the deviation from planarity of the faces of the constructed object were minimized. This enhanced principle took care of the

[7] Unfortunately, their article does not tell us how.

counterexamples cited.

For use in the present context, we can simplify the MSDA algorithm by requiring it to search only over the space of parallelepipeds. This simplified MSDA algorithm will work fine for conforming p-images, as expected. It will not give satisfactory answers, however, in the case of the non-conforming ones. The reason is that the algorithm can find a solution with angles close to ninety degrees (thereby minimizing the standard deviation of angles) by moving the z-coordinates of certain points to extreme depths. Such z-values, however, are quite unrealistic as visual interpretations.

By the same token, the Leclerc and Fischler enhanced algorithm will also fail for non-conforming p-images. We know this because, by constraining our MSDA algorithm to search only over the space of parallelepipeds, we already guarantee that the faces of the constructed objects will be planar (and we know that the algorithm fails for this case). Therefore, the Leclerc and Fischler enhancement that minimizes the deviation from planarity cannot help us for non-conforming p-images.

11 Summary.

We have provided a complete theory of the three-dimensional interpretation of a class of line-drawings called p-images, a subset of parallelogram meshes. Despite the simplicity of p-images, their interpretation has not hitherto been handled satisfactorily in the vision literature.

Specific questions answered by the theory are the following: What are the dimensions and pose of the perceived objects? Why are some p-images seen as rectangular solids, while others are seen as skewed, even though there is no obvious distinction between the images? Why are p-images seen as three-dimensional objects? When p-images are rotated in three dimensions, why are the image-sequences perceived as distorting objects—even though structure-from-motion would predict that rigid objects would be seen? Why are some three-dimensional parallelepipeds seen as radically different when viewed from different viewpoints?

We have also discussed the special case that arises when there are right angles in the p-image. This case represents a singularity in the equations and is mystifying from the vision point of view. It would seem that in this case the vision system does not follow the ordinary rules of geometry but operates in accordance with other (and as yet unknown) principles. This puzzle remains unexplained.

Acknowledgments

I would like to thank Boris Katz, Rodney Brooks and Berthold Horn, who read an early draft and provided insightful comments. I very much appreciate their help.

Appendix A. The theorem of Section 3.

We show that every planar triple vertex V is the (orthographic) image of some right-angled triple vertex, unless V contains a right angle or an odd number of acute angles.

Consider a triple vertex in space (Figure 9(b)). We write the vectors from the central point to the extremities.

(4) $\mathbf{V}_1 = (x_1 - x_0)\mathbf{i} + (y_1 - y_0)\mathbf{j} + (z_1 - z_0)\mathbf{k}$

(5) $\mathbf{V}_2 = (x_2 - x_0)\mathbf{i} + (y_2 - y_0)\mathbf{j} + (z_2 - z_0)\mathbf{k}$

(6) $\mathbf{V}_3 = (x_3 - x_0)\mathbf{i} + (y_3 - y_0)\mathbf{j} + (z_3 - z_0)\mathbf{k}$

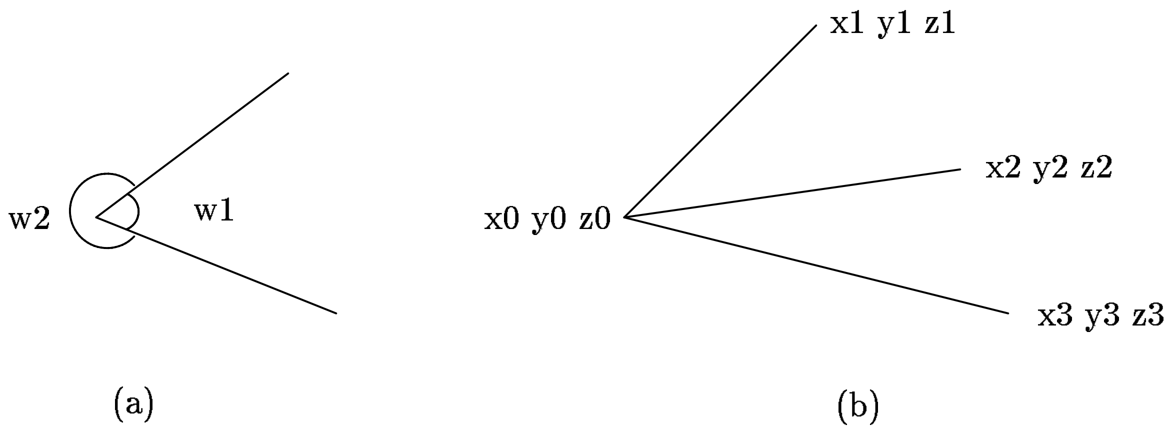

Figure 9: (a) Angle. (b) Triple vertex.

The orthographic images of the three vectors are given by

(7) $\mathbf{v}_1 = (x_1 - x_0)\mathbf{i} + (y_1 - y_0)\mathbf{j}$

(8) $\mathbf{v}_2 = (x_2 - x_0)\mathbf{i} + (y_2 - y_0)\mathbf{j}$

(9) $\mathbf{v}_3 = (x_3 - x_0)\mathbf{i} + (y_3 - y_0)\mathbf{j}$

The three angles of the triple vertex are right angles if the dot-products of the

3D vectors are zero, that is, if

$$(10)\ A_{12} = (x_1 - x_0)(x_2 - x_0) + (y_1 - y_0)(y_2 - y_0) + (z_1 - z_0)(z_2 - z_0) = 0$$

$$(11)\ A_{13} = (x_1 - x_0)(x_3 - x_0) + (y_1 - y_0)(y_3 - y_0) + (z_1 - z_0)(z_3 - z_0) = 0$$

$$(12)\ A_{23} = (x_2 - x_0)(x_3 - x_0) + (y_2 - y_0)(y_3 - y_0) + (z_2 - z_0)(z_3 - z_0) = 0$$

It is also useful to write out the dot-products of the image vectors:

$$(13)\ a_{12} = (x_1 - x_0)(x_2 - x_0) + (y_1 - y_0)(y_2 - y_0)$$

$$(14)\ a_{13} = (x_1 - x_0)(x_3 - x_0) + (y_1 - y_0)(y_3 - y_0)$$

$$(15)\ a_{23} = (x_2 - x_0)(x_3 - x_0) + (y_2 - y_0)(y_3 - y_0)$$

Equations (10), (11) and (12) can then be rewritten:

$$(16)\ (z_1 - z_0)(z_2 - z_0) + a_{12} = 0$$

$$(17)\ (z_1 - z_0)(z_3 - z_0) + a_{13} = 0$$

$$(18)\ (z_2 - z_0)(z_3 - z_0) + a_{23} = 0$$

These three equations can be solved simultaneously to yield:

$$(19) \qquad z_1 = z_0 \pm \sqrt{-a_{12}a_{13}/a_{23}}$$

$$(20) \qquad z_2 = z_0 \pm \sqrt{-a_{12}a_{23}/a_{13}}$$

$$(21) \qquad z_3 = z_0 \pm \sqrt{-a_{13}a_{23}/a_{12}}$$

This tells us that every triple vertex is the image of a right-angled triple vertex, so long as equations (19), (20), and (21) have real solutions. This will always be the case unless the following conditions (a) or (b) occur.

(a) The denominator inside the radical is zero. But this occurs only if one of the image angles is a right angle. Hence none of the angles can be right angles.

(b) The quantity inside the radical is negative, which occurs if the number of negative a_{ij} is even. But recall that a_{ij} is the dot product of image vectors. The dot product will be negative only if the angle is obtuse. Thus condition (b) is that the number of obtuse angles is even (or the number of acute angles is odd). This proves the theorem.

It is easy to show that for (19), (20), and (21) to be joint solutions, we must pick the $+$ sign in all three cases or the $-$ sign in all three cases.

Appendix B. The "compromise" heuristic.

It is impossible to interpret non-conforming p-images as rectangular solids; that is, the three angles among the basis vectors (Figure 3) cannot all be right angles. Here we investigate the "compromise" heuristic, which interprets two of the three angles as right angles and compromises as regards the third. The nature of the compromise is to write z_2 and z_3 as a function of z_1 in two different ways and then to select the value of z_1 that minimize the difference between these two ways. We show that that this approach yields the same equations as Appendix A, except for the sign under the radical. (Thus conforming and non-conforming p-images can be interpreted with a single set of equations by the simple expedient of taking the absolute value of the quantity inside the radical in equations (1), (2) and (3).)

Let us arbitrarily set the point $(x_0 y_0 z_0)$ to be at the origin. Then, using the same notation as Appendix A, we write the dot product of the three space vectors as:

(22) $A_{12} = x_1 x_2 + y_1 y_2 + z_1 z_2$

(23) $A_{13} = x_1 x_3 + y_1 y_3 + z_1 z_3$

(24) $A_{23} = x_2 x_3 + y_2 y_3 + z_2 z_3$

Likewise we write the dot products of the image vectors:

(25) $a_{12} = x_1 x_2 + y_1 y_2$

(26) $a_{13} = x_1 x_3 + y_1 y_3$

(27) $a_{23} = x_2 x_3 + y_2 y_3$

We can rewrite (22), (23), and (24) in terms of (25), (26), and (27):

(28) $A_{12} = a_{12} + z_1 z_2$

(29) $A_{13} = a_{13} + z_1 z_3$

(30) $A_{23} = a_{23} + z_2 z_3$

Suppose we let the angles 1-0-3 and 2-0-3 be right angles. Thus, we set equations (29) and (30) to zero.

We now write z_2 and z_3 as a function of z_1, getting:

(31) $z_2 = z_1(a_{23}/a_{13})$

(32) $z_3 = -a_{13}/z_1$

Using different pairs of angles to be right angles, we can get

(33) $z_2' = -a_{12}/z_1$

(34) $z_3' = z_1(a_{23}/a_{12})$

and also

(35) $z_2'' = -a_{12}/z_1$

(36) $z_3'' = -a_{13}/z_1$

We can think of z_2, z_2', and z_2'' as different "estimates" of z_2, and similarly for z_3, z_3', and z_3''. We wish to find the value of z_1 which minimizes the differences between these estimates. We think of that value of z_1 as a good compromise.

Let us set $E1 = z_2 - z_2'$, , $E2 = z_3 - z_3'$, and $D = E1^2 + E2^2$. If we differentiate D with respect to z_1, set the result to 0, and solve for z_1, we get

(37) $z_1 = \pm\sqrt{\pm a_{12}a_{13}/a_{23}}$

Combining (37) with (35) and (36) gives us

(38) $z_2 = \pm\sqrt{\pm a_{12}a_{23}/a_{13}}$

(39) $z_3 = \pm\sqrt{\pm a_{13}a_{23}/a_{12}}$

Equations (37), (38), and (39) are the same as equations (19), (20) and (21) of Appendix A, except for the sign under the radical. Thus a single set of equations will suffice for the interpretations of p-images, both conforming and non-conforming, if we take the absolute value of the quantity under the radical.

If we select the other choices for the two angles to make into right angles, we will again get the same answer.

Appendix C. A three-dimensional skewed parallelepiped.

The object discussed in Section 8 and shown with a rotation angle of 0 in Figure 7 is the following skewed parallelepiped:

(OBJECT

:POINTS ((0.5 -4.0 5.05) (-2.0 -4.0 0.72) (-1.1 -0.17 2.27) (3.6 0.17 2.06) (-3.6 -0.17 -2.06) (2.0 4.0 -0.72) (-0.5 4.0 -5.05) (1.1 0.17 -2.27))

:LINES ((0 1) (0 2) (0 3) (1 7) (3 7) (1 4) (6 7) (3 5) (4 6) (2 5) (5 6) (2 4)))

In this notation the points are implicitly numbered 0...7. Thus the first line, (0 1), connects point 0 to point 1. The center of the object is at the origin.

References.

Barrow, H.G. and Tenenbaum, J.M. (1981) Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence* **17**, 75-116.

Brady, M. and Yuille, A. (1983) An extremum principle for shape from contour. A.I. Memo 711, Artificial Intelligence Laboratory, MIT.

Clowes, M.B. (1971) On seeing things. *Artificial Intelligence* **2**(1) 79-116.

Friedberg, S.A. (1986). Finding axes of skewed symmetry. *Computer Vision, Graphics, and Image Processing* **34** 138-155.

Horn, B.K.P., (1986) *Robot Vision* (MIT Press, Cambridge, Mass.).

Huffman, D.A. (1970) Impossible objects as nonsense sentences, in: Meltzer, B. and Michie, D., Eds., *Machine Intelligence* **6**, (Edinburgh University Press, Edinburgh).

Kanade, T. (1981) Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence* **17** 409-460.

Leclerc, Y.G., and Fischler, M.A. (1992) An optimization-based approach to the interpretation of single line-drawings as 3D wire-frames. *International Journal of Computer Vision* 9:2, 113-136.

Mackworth, A.K. (1974) Interpreting pictures of polyhedral scenes. *Artificial Intelligence* **4**(2), 121-137.

Marill, T. (1991) Emulating the human interpretation of line-drawings as three-dimensional objects. *International Journal of Computer Vision* 6:2, 147-161.

Marill, T. (1992) Why do we see three-dimensional objects? A.I. Memo 1366, Artificial Intelligence Laboratory, Massachusetts Institute of Technology.

Ullman, S. (1979) *The Interpretation of Visual Motion* (MIT Press, Cambridge, Mass.).