

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Working Paper 320

February 1989

**Further Evidence Against the Recovery Theory  
of Vision**

**Thomas Marill**

**Abstract**

The problem of three-dimensional vision is generally formulated as the problem of recovering the three-dimensional scene that caused the image.

We have previously presented a certain line-drawing and shown that it has the following property: the three-dimensional object we see when we look at this line-drawing does not have the line-drawing as its image. It would therefore be impossible for the seen object to be the cause of the image. Such an occurrence constitutes a counterexample to the theory that vision recovers the scene that caused the image.

Here we show that such a counterexample is not an isolated case, but is the rule rather than the exception. Thus, as a general matter, the three-dimensional scenes we see when we look at line-drawings do not have these drawings as their image. This represents further evidence against the recovery theory.

A.I. Laboratory Working Papers are produced for internal circulation, and may contain information that is, for example, too preliminary or too detailed for formal publication. It is not intended that they should be considered papers to which reference can be made in the literature.



## 1. Introduction.

We have previously shown [1] an example of a line-drawing having the following property: the three-dimensional scene (or object) that we see when we look at this particular line-drawing does not have the drawing as its image. Expressed in terms of the retinal image, the situation is as follows: let  $R$  be the retinal image that obtains when we look at the drawing; then what we see when  $R$  obtains is something that could not have caused  $R$ .

This occurrence represents a counterexample to the “recovery” theory of vision, i.e., the theory that the vision process, given a certain image as input, recovers the object or scene that caused the image.

In the present paper we expand this result to show that this counterexample is not a special case or an “illusion”, but represents the usual state of affairs.

## 2. Informal discussion of result.

In this section, we give an informal overview of the result. The remainder of the paper represents a more formal approach to the same material.

(1) We are all familiar with the fact that, in general, when a three-dimensional object moves, its image changes shape.

For example, we can look at Fig. 1, which shows the perspective image of a cube being translated in the  $x$ -direction. As the cube moves, its image not merely translates but actually deforms.

(2) We are also familiar with the fact that when a drawing or painting moves with respect to the line of sight, the perceived three dimensional objects do not change shape. (If this were not the case, as we walked past a painting in a museum, the perceived objects would deform).

We can demonstrate this fact to ourselves by means of Fig. 2. If we look at point  $D$ , we see a cube. If we now look at point  $E$  (or any other point), we again see the same cube; the cube has not changed shape, but has simply undergone translation with respect to the line of sight. (To be more precise about this, we should Fig. 2 perpendicular to the line of sight and then translate it so that point  $D$ , point  $E$ , or whatever point we choose, lies on the line of sight. However, it really doesn't matter whether we do this or not, since the phenomenon is quite insensitive to the details of viewing condition.)

(3) We see, therefore, that if a three-dimensional object undergoes translation its image changes shape. This is the same as saying that if an image does not change shape, it cannot be the image of a three-dimensional object undergoing translation.

Obviously, a line-drawing does not change shape under translation. Therefore it cannot be the image of three-dimensional object being translated.

However, when we look at a line-drawing being translated, we see a three-dimensional object which does not change shape, but merely undergoes translation.

It follows therefore that when a line-drawing is translated, the object that we see cannot have that line-drawing as its image.

That being the case, the object could not be the cause of the image, and it is not meaningful to think of that object as having been “recovered” by the vision process.

### 3. The concept of the extension of an image.

It will simplify our discussion to define the concept of the extension of an image.

By the *extension of an image* I, we mean the set of all three-dimensional objects or scenes that have I as their image, i.e., as their perspective projection.

When we speak of the image of an object, we understand this image to be defined with respect to a center of projection and image plane. Likewise, when we speak of an extension, we understand this extension to be defined with respect to a center of projection and image plane.

If I is a line-drawing, the extension has an infinite number of members.

In terms of this concept, then, we can say that our previous counterexample [1] shows that, for a certain line-drawing D, the object that is perceived when we look at D is not in the extension of D.

What we show in the present paper is that, in general, when we look at any line-drawing, the seen three-dimensional object, if any, is not in the extension of that line-drawing.

### 4. Further definitions.

It will be useful to introduce some further definitions.

A *wire-frame object* is a set of fixed straight-line segments in space.

A *non-trivial wire-frame object* is a wire-frame object having at least one line-segment not parallel to the image plane.

A *line-drawing* is a set of fixed coplanar straight-line segments.

A *non-trivial line-drawing* is a line-drawing which, when viewed under the standard viewing condition, causes the viewer to see a non-trivial wire-frame object (e.g., Fig. 2).

The *standard viewing condition* is defined as follows. We place the eye (strictly speaking, the optical center of the eye) at the origin of a left-handed coordinate system with the line of sight along the positive  $z$ -axis. The line-drawing is placed in the image plane  $z=f$ ; the points in the line-drawing are defined in terms of their  $(x,y)$ -coordinates in the image plane. For the sake of specificity and to have a comfortable viewing distance we let  $f = 24$  inches.

## 5. Two theorems.

**Vanishing point theorem.** Let the center of projection be at the origin of a left-handed coordinate system and the image plane be  $z = f$ . Let  $L$  be a line in space parallel to the vector  $\mathbf{v} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$  ( $C > 0$ ). Then the perspective image of  $L$  is a half-line in the image plane terminating at the "vanishing point"  $(Af/C, Bf/C)$ .

*Proof.* Since we have  $C > 0$ , we are dealing with lines that are not parallel to the image plane. Such lines can be represented by their direction (parallel to the vector  $\mathbf{v}$ ) and their intersection  $(x_0, y_0)$  with the  $x$ - $y$  plane.

Such a line is the path traced out by the point  $P : (x_0 + tA, y_0 + tB, tC)$  as  $t$  goes from minus to plus infinity. However, the perspective projection of this point is defined only when the  $z$ -coordinate of  $P$  is positive, i.e., when  $t > 0$ .

The image (perspective projection) of this line is then the path traced out by the point

$$P_p : (f(x_0 + tA)/tC, f(y_0 + tB)/tC)$$

i.e., by the point

$$P_p : (fx_0/tC + fA/C, fy_0/tC + fB/C)$$

as  $t$  goes from 0 to plus infinity.

But as  $t$  goes to plus infinity, the point  $P_p$  goes to  $(Af/C, Bf/C)$ . Q.E.D.

**Extension theorem.** Let  $I$  be a line-segment in the image plane and  $I^*$  be a

non-collinear translation of  $I$  in that plane. Let  $L$  be any line-segment, not parallel to the image plane, in the extension of  $I$ . Then there is no line segment  $L^*$  in the extension of  $I^*$  such that  $L^*$  is a translation of  $L$ .

**Proof.** Assume the negation of the theorem; i.e., assume that under the conditions of the theorem there is a line segment  $L^*$  in the extension of  $I^*$  such that  $L^*$  is a translation of  $L$ .

Then  $L$  and  $L^*$  are parallel or collinear. Either way, by the vanishing point theorem (above), they have the same vanishing point in the image plane.

By hypothesis,  $I$  and  $I^*$  are not collinear and are the images of  $L$  and  $L^*$  respectively. Therefore  $I$  and  $I^*$  must lie on lines that intersect at their common vanishing point.

But  $I^*$  is a non-collinear translation of  $I$ ; therefore  $I^*$  is parallel to  $I$ . Hence  $I^*$  and  $I$  cannot lie on lines that intersect.

We have derived a contradiction, so the negation of the theorem must be false. Q.E.D.

## 6. The central argument.

Recovery theory holds that vision recovers the object (or scene) that causes the image. For the case of human vision the theory implies that the perceived object is in the extension of the image. We now demonstrate that this is not generally the case when the image is a line-drawing.

(1) Assume the recovery theory is true; that is, assume the perceived object is in the extension of the image.

(2) Consider any non-trivial line-drawing  $D$ . By hypothesis, the object  $O$  which is seen when we look at  $D$  under the standard viewing conditions is in the extension of  $D$ .

(3) Since  $D$  is non-trivial,  $O$  has a line  $L$  not parallel to the image plane. Let its image in  $D$  be  $l$ .

(4) Let  $D^*$  be a line-drawing which is a translation of  $D$  in a direction not parallel to  $l$ . The object  $O^*$  which is seen when we look at  $D^*$  is in the extension of  $D^*$ .

(5) Let  $l^*$  be the line-segment in  $D^*$  which is the translation of  $l$  in  $D$ . And let  $L^*$  be a line-segment in  $O^*$  which has  $l^*$  as its image.

(6) (At this point we have  $l$  a line-segment in the image plane and  $l^*$  its non-collinear translation in that plane; and we have  $L$ , a line-segment not parallel to

the image plane, in the extension of  $l$ .) By the extension theorem, then, there is no line-segment in the extension of  $l^*$  that is a translation of  $L$ . Since  $L^*$  is in the extension of  $l^*$ , it cannot be a translation of  $L$ . And since object  $O^*$  contains  $L^*$  and object  $O$  contains  $L$ ,  $O^*$  cannot be a translation of  $O$ .

(7) Thus we conclude that in general,<sup>1</sup> in going from  $O$  to  $O^*$ , the object cannot merely translate but must rotate or change shape.

(8) But the conclusion in step (7) is clearly incorrect. When we translate a line-drawing, the objects we see do not rotate or change shape, but merely undergo translation. This is a fact of everyday experience that can readily be demonstrated.

(9) In arriving at our incorrect conclusion, the only fact of vision that was used was the assumption that the seen objects are in the extension of the drawings. All other facts were facts of geometry. One can therefore conclude that the assumption is false.

## 7. Conclusions.

We have shown that in general the three-dimensional objects we see when we look at line-drawings are not in the extensions of these drawings. The objects, therefore, could not be the cause of the images. This result represents further evidence against the recovery theory.

## Reference.

1. T. Marill, "A counterexample to the theory that vision recovers three-dimensional scenes," A.I. Working Paper 319, MIT A.I. Lab, November 1988.

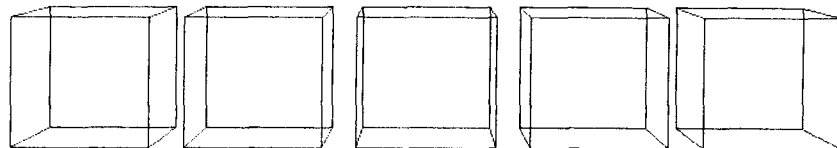
## Figure legends.

1. Perspective projection of a cube being translated in x-dimension.
2. A line-drawing.

---

<sup>1</sup>Strictly speaking, it is necessary that there be a line  $l$  in  $D$ , not parallel to the translation vector, where  $l$  is the image of a line  $L$  not parallel to the image plane. It is possible for special situations to arise in which this condition does not hold. However, these special situations will be rare and in general the condition will hold.

```
(generate-picture-list-and-draw-perspective 0 -5 40 0 0 0)  
NIL  
(generate-picture-list-and-draw-perspective -5 -5 40 0 0 0)  
NIL  
(generate-picture-list-and-draw-perspective -10 -5 40 0 0 0)  
NIL  
(generate-picture-list-and-draw-perspective 5 -5 40 0 0 0)  
NIL  
(generate-picture-list-and-draw-perspective 10 -5 40 0 0 0)  
NIL
```



Lisp Listener 1

[Fri 10 Feb 11:40:27] Mar 11

CL-USEP:

Fig 1

Lisp Machine Worker 4

FIGURE 1



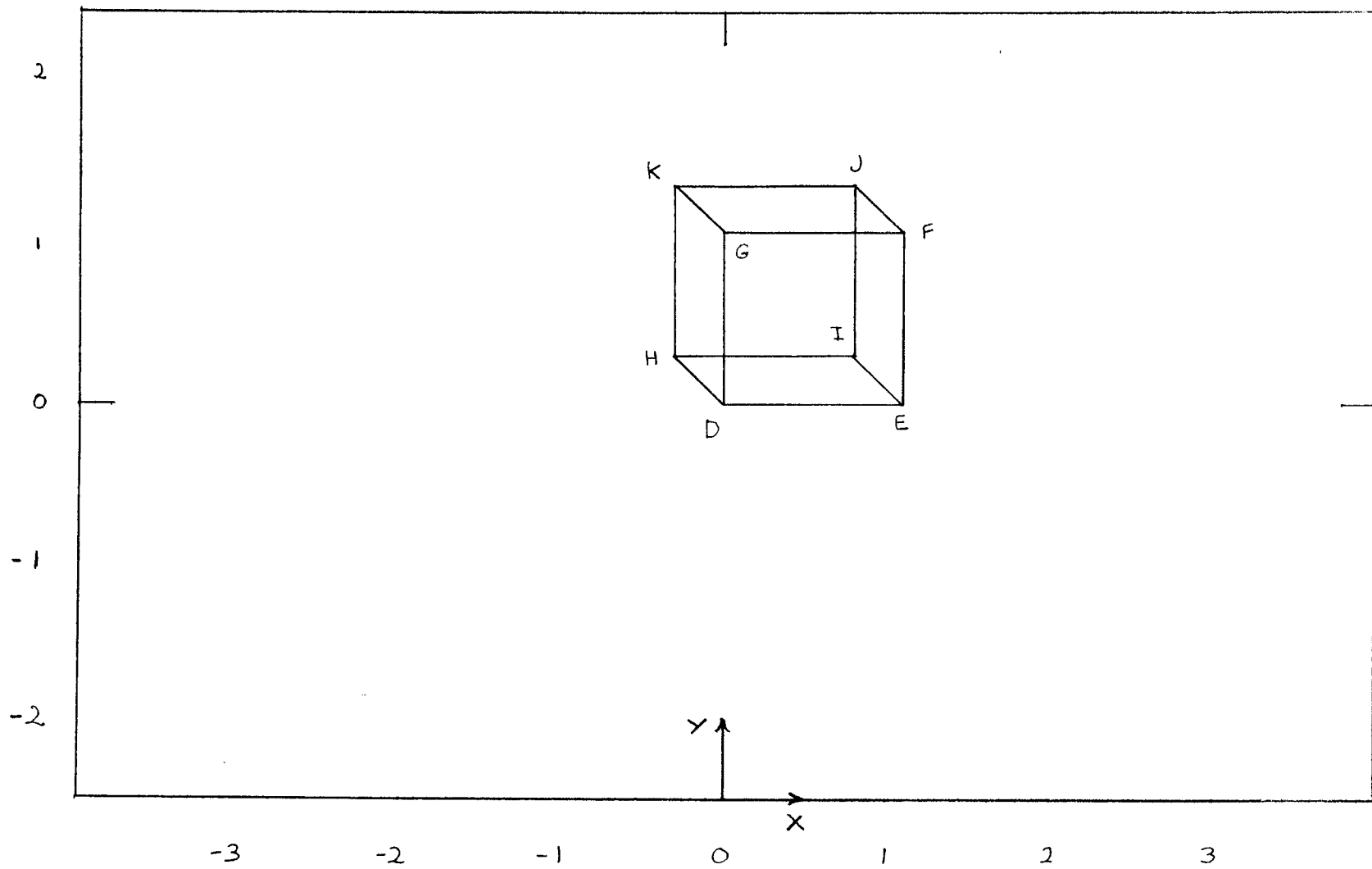


FIGURE 2