ICOT Technical Report: TR-231

TR-231

Cognitive Model for Quantitative Interpretation of Line Drawings

by

N. Oka

February, 1987

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

Cognitive Model for Quantitative Interpretation of Line Drawings

Natsuki Oka

ICOT Research Center
Institute for New Generation Computer Technology
4-28, Mita 1-chome, Minato-ku, Tokyo 108 JAPAN
csnet: oka% icot.jp@csnet-relay
uucp: { inria,kddlab,mit-eddie,ukc} !icot!oka

Abstract

The three-dimensional shape of an object can be perceived quantitatively from a single two-dimensional line drawing. The new inclusive cognitive model proposed in this paper explains this type of interpretation for any kind of object. It interprets a drawing as one of many geometrically possible objects which is most likely to produce the drawing. The likelihood of producing the drawing depends on both the likelihood of an object (object likelihood) and the likelihood of a view of the object (view likelihood). Previously published interpretation methods mostly deal with object likelihood. Thus, these methods are applicable only to regular-shaped or familiar objects. It is shown that in the quantitative interpretation of irregular-shaped and unfamiliar objects, view likelihood is important. For practical use, the likelihood of the projected area is proposed as an approximation to view likelihood.

This paper also examines line drawings which communicate the three-dimensional shape of objects. In this situation, the probability that an interpretation based on the view likelihood is correct increases. This paper proposes a heuristic for the interpretation of communication-oriented drawings, and shows how to make drawings for communication.

TOPIC: Perception and Signal Understanding

SUBTOPIC: Vision

SECONDARY TOPIC: Cognitive Modeling

SCIENCE/ENGINEERING: Science

Cognitive Model for
Quantitative Interpretation of Line Drawings

Natsuki Oka

ICOT Research Center
Institute for New Generation Computer Technology
4-28, Mita 1-chome, Minato-ku, Tokyo 108 JAPAN
csnet: oka% icot.jp@csnet-relay
uucp: { inria,kddlab,mit-eddie,ukc} !icot!oka

1 Introduction

Geometrically, the three-dimensional shape of an object cannot be decided quantitatively from a single two-dimensional line drawing such as Fig.1. However, the three dimensional-shape is usually perceived quantitatively from a single view. Hence, drawings must be interpreted based on certain assumptions. The aim of this paper is to clarify these assumptions.

This paper deals with line drawings obtained by orthogonal projection. First, **section 2** examines previously published interpretation methods and points out their limitations. **Section 3** gives an outline of a cognitive model which overcomes the limitations and can interpret any kind of drawing in the same way as human vision. **Section 4** describes the interpretation of communication-oriented drawings. **Section 4.1** deals with drawings obtained by oblique projection.

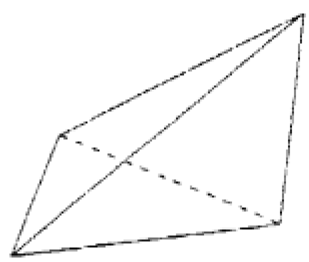


Fig. 1

2 Related Work

Roberts' system [Rob65] determined the position and size of objects using a model which represents each possible type of object in a three-dimensional coordinate system, assuming that each object must be supported in some way. Naturally, this system cannot interpret an object without a model.

The PICAX system [LHR78] handles drawings in which three projected axes are specified. It interprets drawings on the assumption that a line parallel to one of the projected axes represents a line really parallel to one of the real axes.

If it cannot decide the 3-D coordinate of some of the vertices, it asks the user for construction lines for the unknown vertices. A construction line is parallel to the projected z-axis. For example, the user must designate a construction line for the top vertex of a pyramid (Fig.2).

Hirotani et al.'s system [HFO81] also handles the drawings in which three projected axes are specified. It interprets drawings on two assumptions. One is that a line parallel to one of the projected axes represents a line really parallel to one of the real axes. The other is that objects are set on the xy-plane (Fig.3). However, it cannot interpret a pyramid or an oblique prism.

Kanade [Kan81] proposed two heuristics for the interpretation of line drawings: the parallel-line heuristic and the skewed-symmetry heuristic.

The parallel-line heuristic is

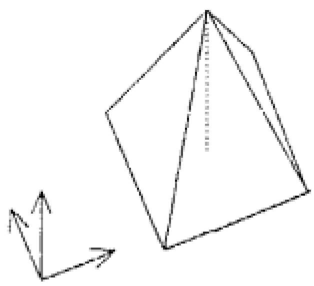


Fig.2

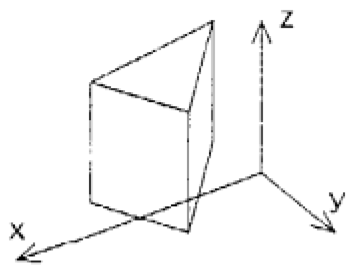


Fig.3

"if two lines are parallel in the picture, they depict parallel lines in the scene."

The skewed-symmetry heuristic is

"skewed symmetry depicts a real symmetry viewed from some (unknown) view direction."

The concept of skewed symmetry is a class of 2-D shapes in which the symmetry is found along lines not necessarily perpendicular to the symmetry axis.

With these heuristics, a drawing such as Fig.4(a) can be interpreted quantitatively as a rectangular prism. However, a drawing such as Fig.4(b) cannot be interpreted quantitatively. The only definite statement that can be made with these heuristics is that it cannot be a right-angled block.

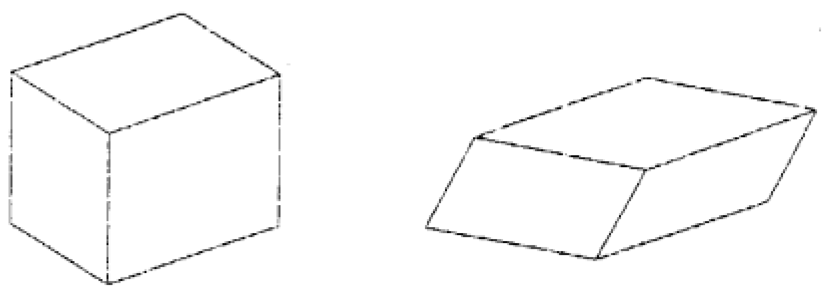
3 Cognitive Model for Quantitative Interpretation

It is common for humans to perceive a three-dimensional shape even when the object includes no symmetrical shape, such as Fig.1 or Fig.4(b). The new cognitive model proposed in this paper can interpret not only symmetrical, but also asymmetrical objects without additional information, such as a construction line.

3.1 The Likelihood Model

The notation $< object, view >$ represents an *object* with a *view*.

The basic idea of the model proposed here is to



(a) (b)

Fig. 4

"interpret a drawing according to the likelihood of $< object, view >$ to the drawing."

This means that humans normally interpret a drawing as the $< object, view >$ which is most likely to produce the drawing, and that it is difficult to perceive an $< object, view >$ with a low likelihood although it can be perceived intentionally. This is called the likelihood model. This type of viewpoint is known as maximum likelihood estimation in statistics.

The likelihood depends on three factors. That is,

$$Lt(drawing; object, view)$$
$$= Lg(drawing; object, view) \times Lo(object) \times Lv(view \mid object)$$

where Lt is the total likelihood, Lg is the geometrical likelihood which represents the geometrical constraints, Lo is the object likelihood which represents the likelihood of an *object* being drawn, and Lv is the view likelihood which represents the likelihood of a *view* being selected on condition that an *object* is drawn. Putting it in another way, Lg corresponds to a bottom up (data-driven) process and Lo and Lv to a top down (model-driven) process.

For example, suppose that drawing D can be interpreted geometrically as object $O1$ with view $V1$ or as object $O2$ with view $V2$ such that

$$Lg(D; O1, V1) = Lg(D; O2, V2).$$

If

$$Lo(O1) = Lo(O2),$$
$$Lv(V1 \mid O1) > Lv(V2 \mid O2)$$

hold, then

$$Lt(D; O1, V1) > Lt(D; O2, V2)$$

holds. This means that object $O1$ is normally perceived from the drawing.

It seems that humans learn how to reconstruct three-dimensional shapes from two-dimensional images from many examples of image-solid pairs. The likelihood model seems to be a natural result of this type of learning.

3.2 Geometrical Likelihood

Since human judgment about geometrical possibility is not very strict, the value of the geometrical likelihood becomes maximum at strictly possible objects and decreases continuously around the maximum.

Although the likelihood model can handle ambiguities in geometrical possibility judgment of human vision, it is not the main aim of this paper to treat such ambiguities. So for simplicity, this paper regards geometrical likelihood as a two-valued function such that

$$Lg(D;O,V) = \begin{cases} 1 & \text{if } <O,V> \text{ is geometrically strictly possible} \\ 0 & \text{otherwise.} \end{cases}$$

The paper does not refer to the mechanism for judging geometrical possibility. The problem treated in this paper is

"How do we choose one from many kinds of geometrically possible objects?"

3.3 Object Likelihood

Object likelihood is the likelihood of an object being drawn. This likelihood is basically equal among humans, but it varies reflecting the interpreter's individual visual experience and individual expectations. That is, it depends on what he has seen and what he expects to see. This variation naturally corresponds to variations in interpretation among humans and variations among context.

In previous work, drawings have been interpreted assuming that objects are similar to the known model [Rob65], or that objects have a symmetric shape [Kan81], or that objects have parallel lines [Kan81]. These assumptions correspond to restrictions by the simplified object likelihood function. That is, in previous work the object likelihood is a two-valued function such that

$$Lo(O) = \begin{cases} 1 & \text{if O is an assumed object} \\ 0 & \text{otherwise.} \end{cases}$$

By these methods, complete interpretation is possible if and only if exactly one object is geometrically possible among assumed objects. Hence, the drawings which can be interpreted by these methods are strictly limited.

For the interpretation of general drawings, a many-valued object likelihood function is necessary as well as the view likelihood function. Examples of this case are shown in **section 3.5**.

A useful, new heuristic which is another example of restriction by the two-valued object likelihood function is proposed here, with an example of its application. This heuristic is somewhat analogous to the skewed-symmetry heuristic, and is called a whole-symmetry heuristic.

The whole-symmetry heuristic is to

"interpret a drawing assuming it is a symmetrical object if possible."

In Fig.5, even if it is assumed that the legs and the top plane of the table are rectangular parallelepipeds and that the top plane is supported by legs, the relative position of the legs to the top plane cannot be determined geometrically because part of the legs are hidden. With the whole-symmetry heuristic, the relative position can be determined quantitatively in the same way as human interpretation. If the table is asymmetric, the interpretation becomes contradictory. Hence, it can be concluded that it is not symmetric.

3.4 View Likelihood

View likelihood is the likelihood of a view being selected on condition that an object is drawn. For example, Fig.6 shows various views of a skewed parallelepiped. It is clear that the likelihood of view (a) is large, that that of view (b) is small, and that that of

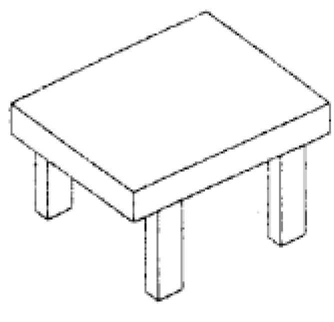


Fig.5

view (c) is smaller. It is affected by context and varies slightly individually.

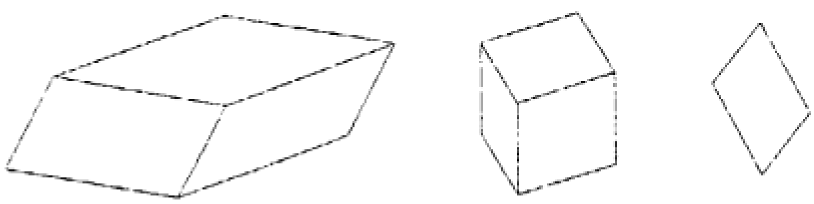
Some researches can be regarded as related to view likelihood. The assumption that objects are supported [Rob65], [HFO81] is one particular case of view likelihood. The parallel-line heuristic [Kan81] can also be regarded as an example of view likelihood. Kanade [Kan81] states that the least slanted planes are the most reasonable selection among the surface orientations which satisfy the skewed-symmetry heuristic if no other constraints are available. This is another specific example of view likelihood.

As shown above, view likelihood has been considered only in particular cases or only for a plane. This paper proposes a new approximation to view likelihood so that the concept of view likelihood can be applied to any objects in general cases. This is the projected area likelihood. Although the projected area is insufficient to represent a view and view likelihood can also be affected by context or individual experience, the projected area likelihood is a good approximation because it reflects the total figure and can be calculated easily.

The larger the projected area, the larger its likelihood. The projected area likelihood is at a maximum at the view angle at the maximum projected area. This can be understood intuitively if you think of a line or a plane as the extreme case of an object.

This paper proposes a new heuristic which corresponds to the two-valued approximate view likelihood function such that

$$Lv(V \mid O) = \begin{cases} 1 & \text{if the maximum likelihood view of } O \text{ is } V \\ 0 & \text{otherwise.} \end{cases}$$



(a) (b) (c)

Fig. 6

It is called a projected-area heuristic. This heuristic can be applied to any kind of object and is especially useful for unfamiliar objects which include no symmetrical shape. As is shown in a later section, the probability that interpretation based on view likelihood is correct increases for communication-oriented drawings. Thus, the probability that interpretation based on this heuristic is correct also increases in that situation.

The projected-area heuristic is to

> "interpret a drawing assuming that it is projected from the angle from which the projected area becomes maximum."

This paper shows two examples of the application of this heuristic.

Fig.7 shows a general triangular pyramid. The projected-area heuristic states that this is a view of the maximum projected area. From the assumption of a local maximum, it can be shown that

$$z(A) = z(C),$$
$$z(B) = z(D)$$

where the projection direction is parallel to z-axis, and $z(A)$ represents the z coordinates of apex A.

From the assumption of the total maximum, it can be shown that

$$\mid z(A) - z(B) \mid \leq limit$$

where $limit$ is a value which can be calculated from the drawing.

Suppose that Fig.8 shows a parallelepiped. According to the projected-area heuristic, it can similarly be shown that

$$z(A) = z(C) = z(H),$$
$$z(B) = z(E) = z(G),$$
$$\mid z(A) - z(B) \mid \leq limit.$$

This paper has shown what can be determined by the projected-area heuristic. There remains some ambiguity in interpretation. A full interpretation can be obtained using both object likelihood and view likelihood, as shown in the following section.

3.5 Total Function of the Likelihood Model

This paper has described the function of object likelihood and view likelihood individually. This section explains the total function of the likelihood model.

First, the relation between object likelihood and view likelihood is clarified. If there is one object with an outstandingly high object likelihood among geometrically possible objects corresponding to a drawing, that drawing is interpreted as that object, regardless of the view likelihood, because the variation of view likelihood is generally not very large. Otherwise, view likelihood is important. The product of object likelihood and view likelihood determines the interpretation.

Previous work mostly treats the former case. **Example 1** is an example of the former case. **Examples 2** and **3** are examples of the latter case. These two examples cannot be interpreted by conventional methods.

A. Example 1 (rectangular parallelepiped)

Geometrically, the object shown in Fig.9 can be various hexahedrons. However, object likelihoods of rectangular parallelepipeds are generally much higher than those of other hexahedrons, and there is one geometrically possible rectangular parallelepiped. Thus, view likelihood does not affect the interpretation and the rectangular parallelepiped is perceived.

B. Example 2 (skewed parallelepiped)

Geometrically, the object shown in Fig.10 can be various hexahedrons other than rect-

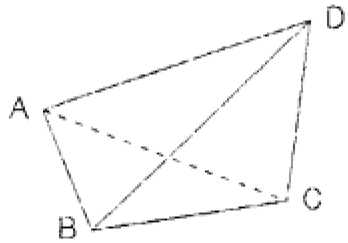


Fig.7

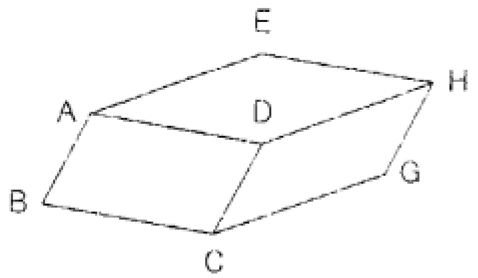


Fig.8

angular parallelepipeds. Object likelihoods of skewed parallelepipeds are much higher than those of other possible hexahedrons. However, there are many possible skewed parallelepipeds, each of which has only a slightly different object likelihood. The view likelihood does not vary very much, either. Thus, both the object likelihood and the view likelihood affect the maximum total likelihood, and the skewed parallelepiped with the maximum total likelihood is perceived. By approximation, of objects supported by the projected-area heuristic, the one with the maximum object likelihood is perceived.

C. Example 3 (triangular pyramid)

In Fig.11, the object likelihood of possible objects does not vary greatly. Thus, in the same way as **example 2**, both the object likelihood and the view likelihood affect the maximum total likelihood. The pyramid with the maximum total likelihood is perceived. By approximation, of objects supported by the projected-area heuristic, the one with the maximum object likelihood is perceived.

4 Line Drawings for Communication

The more important the view likelihood, the less accurate the interpretation by the likelihood model is likely to be. However, this is also the case for human vision. The likelihood model simulates human vision well from this point of view, too.

If drawings are limited to those which are drawn to communicate the three-dimensional shape of an object, the accuracy of interpretation by the likelihood model increases

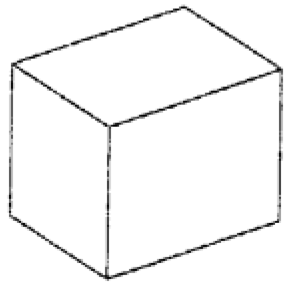


Fig. 9

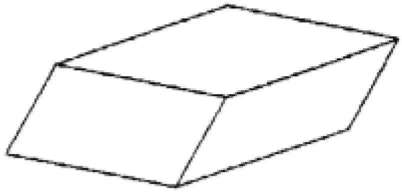


Fig. 10

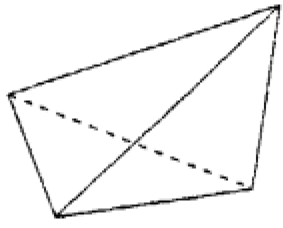


Fig. 11

greatly. This is also true in human vision. This is because drawings are made to communicate the three-dimensional shape as accurately as possible taking into account usual human interpretation.

It is useful to point out here that views with high likelihood correspond to views which communicate a great deal of information. This is easy to understand if the projected area is regarded as a measure of the amount of information. That is, views with a high likelihood have a large projected area, therefore they have a large amount of information.

4.1 Objects which Include an Irregular Element

This section considers objects which include an irregular element. For example, the irregular element of an oblique prism is the slant of the axis, and the irregular element of a pyramid with a unique-shaped base is its base. When humans draw objects which include an irregular element such as these, they draw it paying attention to the irregular element. Thus, they often draw the irregular element as parallel to the plane of projection by oblique projection. This means that the likelihood of such a view increases.

Consequently, it is valid to interpret a drawing of an object which includes an irregular element assuming that the irregular element is parallel to the plane of projection if it is drawn for communication by oblique projection.

The following heuristic refers to a special case of discussion above. Thus, it is especially useful for the interpretation of oblique objects, and can be regarded as a kind of view likelihood. It is called an axis-direction heuristic.

The axis-direction heuristic is to

> "interpret a drawing assuming that the axis of the object is parallel to the plane of projection."

It is known that many objects can be represented as a generalized cylinder [AB73]. The axis here means the axis of an object which is represented as a generalized cylinder. One example of application of this heuristic is shown below.

The base of the pyramid in Fig.12(a) can be interpreted quantitatively by assuming the default projection of coordinate axes as in Fig.12(b) and by assuming that base BCD is on the xy-plane. These assumptions are a kind of view likelihood. If the base is symmetric, these assumptions are unnecessary. The position of apex A can be determined using the axis-direction heuristic which states that axis AG is parallel to the plane of projection.

4.2 Principle of Drawing a Picture

Based on the likelihood model, this section shows how to make a drawing to communicate the three-dimensional shapes of objects, including unfamiliar, asymmetrical objects.

The principle is to

"draw the view which has the maximum likelihood."

The following guides are similar to the principle.

"Draw at a familiar view angle."

"Draw so that the projected area is as large as possible."

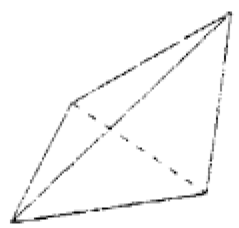
"Draw so that the depth is as small as possible."

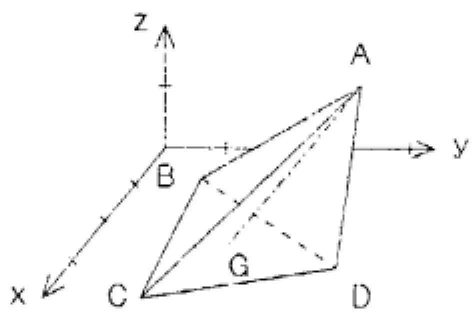
The following guide corresponds to a part of the principle.

"Do not draw so that it can be perceived accidentally as a regular object."

To draw an object which includes an irregular element by oblique projection,

"draw so that the irregular element is parallel to the plane of projection."


(a)



(b)

Fig. 12

5 Conclusion

This paper gave an outline of a new cognitive model which gives a unified explanation of the interpretation of various objects based on object likelihood and view likelihood. Consideration of view likelihood enables interpretation of objects which have no symmetric shape. The projected area likelihood was proposed as a good approximation to view likelihood. It was shown that the probability that an interpretation based on view likelihood is correct increases if drawings are limited to those which are drawn for communication.

This likelihood model can be applied to interpretation for input of CAD systems or animation systems. The present condition is that users of those systems must designate the full information of shapes accurately even when the accurate shape is unnecessary for their purpose, because there is no way to communicate an outline. The likelihood model is suitable for approximate recovery of shapes from line drawings.

It is planned to carry out psychological experiments to determine the likelihood functions and to implement the model from the results of the experiments. The experiments will also test the hypothesis which states that humans usually perceive the object with the maximum likelihood, because it is possible that humans perceive the average shape according to the likelihood. Another direction of future research is to generate the likelihood model automatically by learning from examples.

Acknowledgments

I would like to thank Tsutomu Matsuura for helping with proofs, and Shunichi Amari, Katsuhiko Noda, Yasuo Iwashita and Kouichi Furukawa for giving me research opportunities and encouraging me. Thanks must also go to Fumiyuki Kato and Hirohisa Seki for helpful discussions.

References

[AB73] G.J. Agin and T.O. Binford. Computer Description of Curved Objects. In *Proc. 3rd Int. Joint Conf. Artificial Intelligence*, pages 629–640, 1973.

[HFO81] T. Hirotani, Y. Fukui, and T. Ohira. A Study on Interactive Computer Input System. In *Proc. 3rd Symposium on Computer Use*, pages 343–348, Architecture Society of Japan, 1981. in Japanese.

[Kan81] T. Kanade. Recovery of the Three-Dimensional Shape of an Object from a Single View. *Artificial Intelligence*, 17:409–460, 1981.

[LHR78] M. Liardet, C. Holmes, and D. Rosenthal. *Input to CAD Systems: Two Practical Examples*, Artificial Intelligence and Pattern Recognition in Computer Aided Design, pages 403–427. North-Holland, 1978.

[Rob65] L.G. Roberts. *Machine Perception of Three-Dimensional Solids*, Optical and Electro-Optical Information Processing, pages 159–197. MIT Press, 1965.