

A THEORY OF THE ORIGAMI WORLD

Takeo Kanade *

Department of Information Science
Kyoto University, Kyoto, Japan

1. INTRODUCTION

This paper presents a brief summary of a theory of the Origami world[4], a model for understanding line drawings in terms of plane surfaces and for finding their qualitative 3-D shapes by assigning the labels (+, -, $\nearrow$ or $\searrow$) to each line. The labels signify the physical meaning of the lines (Huffman [2]): the label + stands for a convex edge, - for a concave edge, and $\nearrow$ or $\searrow$ for an occluding edge.

The theory is an extension of the work of Huffman [2] and Mackworth[6], but the feature is that it is surface oriented, that the precompiled knowledge is all contained in the augmented junction dictionary, and that the labeling is all symbolically performed using the filtering procedure not only for a consistency check of junction labels as in Waltz[7] but also in the consistency check of surface orientations in the gradient space.

2. SUMMARY OF THE THEORY OF THE ORIGAMI WORLD

2.1 Surface-oriented Assumption

An important feature of the Origami world is that it is surface oriented. This idea can be best illustrated by Fig. 1. Though it appears perfect, the Huffman-Clowes-Waltz labeling scheme for the trihedral world cannot handle it. The reason for this failure is that the trihedral world essentially assumes solid objects, and thus the picture of a box would need to be "super" perfect, as in Fig. 2, in order for it to be handled.

The assumption concerning the geometrical configurations around a vertex in the Origami world is as follows. The planar surfaces meet edge to edge, no more than three surfaces of different orientations meet at a vertex, and the combination of the three orientations is general in the sense

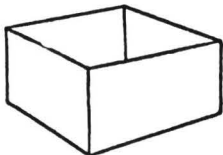


Figure 1. A line drawing of a "box" scene. Though it looks perfect, the trihedral labeling does not work for it.

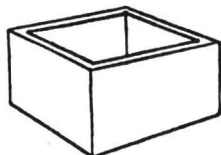


Figure 2. A "super-perfect" line drawing of the "box" scene for the trihedral world.

that they span the three-dimensional space. Let us call such vertices up-to-3-surface vertices.

2.2 Legal Junction Labels for Up-to-3-Surface Vertices

Once we recognize the basic assumptions of the Origami world, we can generate the junction dictionary, as in the Huffman-Clowes-Waltz theory, which contains legal junction labels (i.e., possible line-label combinations) for each junction type. Table 1 gives an idea of the degree of constraints imposed by the up-to-3 surface vertices compared with the trihedral vertices.

2.3 Augmented Junction Dictionary

Necessity: A legal junction label represents a possible configuration of surfaces at a vertex. Consistently labeling a line drawing, so that all the junctions are given legal junction labels, is nothing but checking the consistency of surface interconnections by passing information by means of line labels from one junction to another. The Waltz filtering[7] on junction labels is known to be a good method for doing this. However, the labeling in the Origami world cannot simply rely on the filtering on junction labels. Because of the weaker restriction at the vertices than the trihedral world, a lot of anomalous interpretations exist in which the labeling is consistent, but the whole configuration is not possible[6].

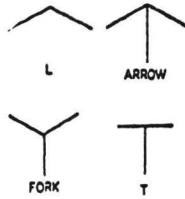
Links: The junction dictionary for the Origami world is augmented for more thorough and global consistency check concerning surface orientations. The surface orientations are conveniently represented by means of the gradient space[2][6]. To each legal junction label is attached the information as to what constraints in the gradient space should be satisfied by the surfaces incident at the junction. Fig. 3 shows two examples. As shown in Fig. 3(a), the constraints are represented by the links which connect a pair of related regions and which include information about the constraints on their gradients. The properties of dual lines of the gradient space are used here.

In such a junction label as shown in Fig. 3(b), which is legal in the Origami world, we have an occluded intersection. It is typically a result of folding a planar surface along BC, thus R1 occludes a part of R2. A little more generally, we can include such cases that the intersection line of R1 and R2 lies within the angle ABC. Therefore, the associated link represents that the gradient G2 of R2 should be inside of the fan-shaped area

* This research was done while the author was visiting Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pa.

Table 1. The sizes of junction dictionaries.

Junction Type	Huffman-Clowes Dictionary	Origami World Dictionary
L	6	8
ARROW	3	15
FORK	3	9
T	4	16



whose origin is at G1, and is bounded by the lines which are perpendicular to AB and BC.

2.4 Labeling Procedure

The labeling procedure of the Origami world uses the augmented junction dictionary. First, the Waltz filtering on junction labels is performed. Then, the procedure begins to assign a junction label, which has been filtered and left, to each junction one by one. When a junction label is assigned to a junction, the constraints represented by the associated links are instantiated by the directions of lines of that particular junction.

Consistency of surface orientations is tested by using these instantiated constraints. The test can be performed by an iterative "filtering" operation defined on a labeled graph called a Surface Connection Graph (SCG), in which a node represents a surface, and an arc represents the instantiated constraint between the surfaces. The feature of this labeling procedure of the Origami world is that all the constraints are maintained symbolically in the SCG during the computation.

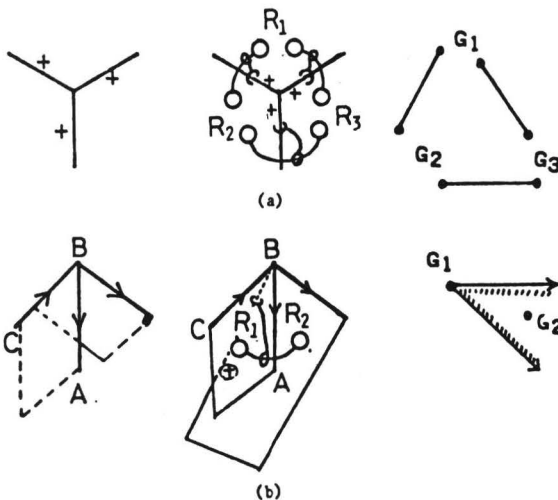


Figure 3. The augmented junction dictionary for the Origami world. Two examples of entries are shown. For each, the first column is the junction label, the second the associated links, and the third the illustration of the relationships of gradients represented by the links.

2.5 The Results of Labeling

As a result of the labeling procedure, we obtain not only a labeled line drawing but also a filtered SCG, which is a summarized description of the constraints on the gradients of the involved surfaces.

Let us see a few examples of labeling. Usually we obtain multiple labelings. The "cube" scene has three labelings (Fig. 4). The "box" line drawing of Fig. 1 has eight labelings, two of which are shown in Fig. 5. The labeling in Fig. 5(a) corresponds to an "ordinary box"; the two front faces form a convex intersection, and partially occlude the rear two faces which form a concave intersection. Fig. 4(b) corresponds to a "squashed box"; the front two faces, as well as the rear two, form a concave intersection.

3. DISCUSSIONS AND COMMENTS

3.1 Origami World and Various Worlds

Assume we have a set of line drawings. We can consider a set of all the combinations in assignments of line labels for those line drawings. A subset exists comprising those interpretations which can be realized by plane surfaces. Let us denote this subset as the Plane Surface World, Spsw. We can also think of a subset, the Consistent Gradient World Scgw, consisting of interpretations in which all the constraints on the gradients of surfaces are completely satisfied. Obviously $Scgw \supset Spsw$.

We can view a labeling procedure as a method consisting of a generator and a tester: given a line drawing, a generator generates interpretations in a certain manner, each of which a tester accepts or rejects based on a certain criterion. Table 2 summarizes various labeling methods [1][2][3][6][7] according to this taxonomy. Various subsets can be defined which are generated by generators or determined as legal by testers. Fig. 6 shows the set inclusion relationships among those subsets. The locations of several example interpretations indicated in it can serve to understand the relationships. (See [4] for details).

3.2 Qualitative Shape Recovery

Note that we have only qualitatively recovered the 3-D shapes by labeling drawings. For example, the labeling of Fig. 4(a) represents a convex corner, but it can be either a cube or a skewed rhomboid. In fact, a "cube" scene (Fig. 4(a)), a "trapezoid-block" scene (Fig. 4(b)) and a "house" scene (Fig. 4(c)) all have the equivalent labelings. In order to recover the shape quantitatively we need to introduce more assumptions together with tools for exploiting them.

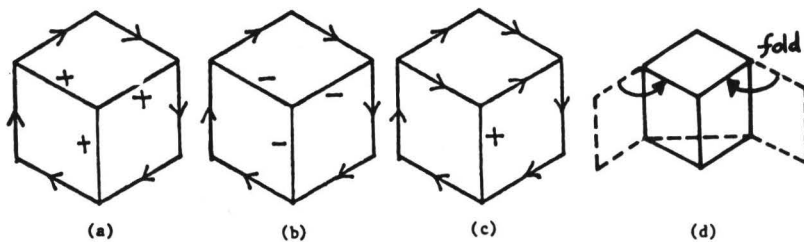


Figure 4. The "cube" scene has three labelings: (a) a convex corner; (b) a concave corner. (c) This represents a shape which can be made as shown in (d)

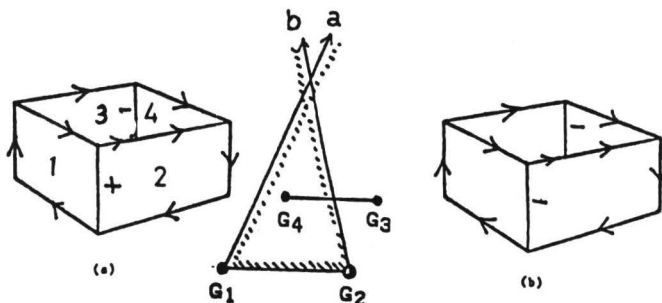


Figure 5. There are eight labelings in total for the "box" scene. Two of them are shown here: (a) corresponds to an ordinary box; (b) a "squashed" box. The diagram represents the constraints in the gradient space for the labeling (a).

3.3 "Natural" and "Unnatural" Interpretations

One of the common impressions we tend to have about those multiple interpretations is that most of them are "unnatural", and therefore it is a disadvantage of the theory to allow them. It is true that we do not usually think of such peculiar shapes as Fig. 4(c) and 5(b) in interpreting these line drawings. It should be very difficult even to imagine those shapes as possible interpretations. However, any labeling is geometrically no more "natural" or "unnatural" than others, unless we use more assumptions.

The above two points (3.2 and 3.3) are actually closely related problems, for whose solution we need to identify more assumptions concerning image formation process and objects in concern. The recent paper [5] by the author presents a theory and a technique which can recover "natural" quantitative 3-D shapes of an object from a single picture, so that we can generate images of the same object as we would see it from other directions.

ACKNOWLEDGEMENT

I thank John Kender, Allen Newell, Raj Reddy, and Steven Shafer for having very stimulating discussions with me.

REFERENCES

- [1] Clowes, M.B., "On Seeing Things", *Artif. Intell.*, 2, p.79, 1971.
- [2] Huffman, D.A., "Impossible Objects as Nonsense Sentences", *Machine Intelligence* 6, Edinburgh University Press, p.295, 1971.
- [3] Huffman, D.A., "Realizable Configuration of Lines in Pictures of Polyhedra", *Machine Intelligence* 8, Edinburgh University Press, p.493, 1977.
- [4] Kanade, T., "A Theory of Origami World", Technical Rep. CMU-CS-144, Carnegie-Mellon University, Pittsburgh, Pa. 1978.
- [5] Kanade, T., "Recovery of Three-Dimensional Shapes of an Object from a Single View", CMU Tech. Report, 1979.
- [6] Mackworth, A.K., "Interpreting Pictures of Polyhedral Scenes", *Artif. Intell.*, 4, p.121, 1973.
- [7] Waltz, D., "Generating Semantic Descriptions from Drawings of Scenes with Shadows", MAC AI-TR-271, MIT., 1972.

Table 2. Various labeling schemes as a method of a generator and a tester.

Method	Generator	Subset	Tester	Subset
Huffman Clowes	Trihedral junction dictionary	S_{tri}		
Waltz	Trihedral junction dictionary with cracks, shadows, etc			
Mackworth	Sequential generation of most connected interpretations		Constructive test on coherence rules in the gradient space	S_{poly}
Huffman			$\phi(\phi')$ point test for all the cut sets in the line drawing	$S_{\phi(\phi')}$
Origami World	Up-to-3-surface junction dictionary	S_{up3}	Filtering of spanning angles on the SCG	$S_{origami}$

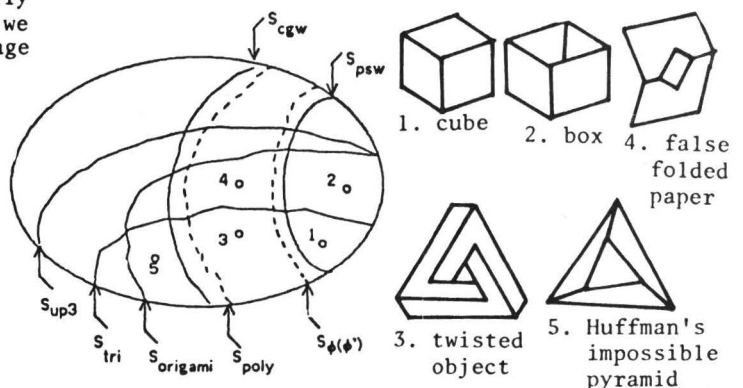


Figure 6. Relationship among various subsets of interpretations.