コンピュータビジョン　56－6
（1988．9．29）

四辺形の視覚的解釈　その2
－重力を受ける長方形の姿勢

徐　剛、　辻　三郎
大阪大学基礎工学部

「四辺形の視覚的解釈　その1」で、四辺形を長方形として解釈することについて述べた。ここでは、重力をうける長方形の姿勢について述べる。重力を受けるものは何かによって支えられなければならない。それは重力の方向に垂直の地平面と知覚されることが多い。カメラと地面との関係、長方形と地面との関係を明らかにし、長方形の向きの拘束条件を得る。

The Visual Interpretation of Quadrilaterals:
The Gravity Regularity

Gang Xu and Saburo Tsuji
Department of Control Engineering, Osaka University

Abstract

We have proposed the rectangularity regularity to be prime constraint in the visual interpretation of quadrilaterals, by which quadrilaterals in image are interpreted as rectangles in space [Xu & Tsuji, 1988]. In this paper we propose another regularity, the gravity regularity. Everything, including the perceiver itself, is attracted by gravity. As a consequence, objects must be supported by something. It is usually perceived to be the ground plane, perpendicular to the direction of gravity, if no evidence indicates otherwise. This paper first analyzes the relation among the camera, the ground and the rectangles supported by the ground, and then derive constraints to determine the rectangle orientation.

1. Introduction

An image is a two-dimensional projection of a three-dimensional scene, and a contour image represents significant changes in surface shape, reflectance and illumination, which are reflected directly or indirectly in the 2-D image. A contour image can be a line drawing, drawn by human hands, which only includes topologically well-defined, semantically significant, but not necessarily positionally accurate contours; or it can be generated by computer from a real image, which may include many noise edges, if not processed very carefully.

Line drawing is probably the most abstract and efficient means of describing our 3-dimensional world in a 2-dimensional manner. Thus it is often used in human communications and is potentially very useful in human-machine interfaces. Although the contours alone do not provide sufficient constraints on the surfaces, humans seem not to have any difficulty in recovering 3-D shapes from the 2-D contours. It is so effortless for our eyes that we rarely pause to ask ourselves how we do. As we try to answer, however, we realize that it is a difficult question [Barnard & Pentland, 1983; Barrow & Tenenbaum, 1981; Brady & Yuille, 1984; Kanade & Kender, 1983; Xu & Tsuji, 1987a,b].

The human visual perception, as a part of the brain, is the product of millions of years of evolution. As a consequence, various regularities of nature have been embedded into the vision system. It is these natural regularities that are secrets of the human vision (and the human perception at large.) They fill in the blank inherent in the mapping from two-dimensionality onto three-dimensionality. Only by understanding them, can we really understand the human vision and further develop any computer vision systems. (See [Pentland, 1986] for a general discussion on the role that natural regularities play in visual perception, and [Ullman, 1979a,b; Reuman & Hoffman, 1986] for how natural regularities play in the visual perception of motion.)

Interpreting an image is different from interpreting a line drawing. Generally images are not taken from purposeful positions and directions; but line drawings are drawn by humans with certain subjective habits or tendency. For example, while it is practically impossible for an image to be projected orthographically, humans tend to assume orthographic projection rather than perspective projection in line drawings. Thus to interpret line drawings it is necessary to take into account how they are drawn and what knowledge, both explicitly and implicitly, takes part in the process. In [Xu & Tsuji, 1988], we have proposed the rectangularity regularity to be prime constraint in the visual interpretation of quadrilaterals, by which quadrilaterals in image are interpreted as rectangles in space. In this paper, we propose the gravity regularity to be an active part in the process, analyze the relations among the perceiver, objects to be perceived and the ground plane that is perpendicular to the gravitational direction and derive constraints to determine rectangle orientations.

2. The Gravity Regularity

Everything, including the perceiver itself, is attracted by gravity. As a consequence, objects must be supported by something. It is usually perceived to be the ground, perpendicular to the direction of gravity, if no evidence indicates otherwise. The gravity regularity is generalized from this universal fact. Unlike the rectangularity, it is a natural regularity objectively existing in the external world. So far it has attracted only a little attention. Kanade *et al.* (1983) analyzes skewed symmetry under gravity. Recently, Sedgwick (1987) reports a production system that generates an interpretation of the environment based on linear perspective information and contact relations between surfaces and the ground. Tsuji *et al.* (1986) also reports a mobile robot that perceives and navigates in an indoor environment with a horizontal flat floor and objects standing vertically on the floor.

To perceive the world is, in essence, to perceive the relations among the perceiver, the ground and the objects on the ground. By introducing the ground, the relation between the perceiver and the rectangles reduces to the sum of the relation between the perceiver and the ground, and the relation between the ground and the rectangles supported by it. All these relations can be described in either a viewer-centered representation or a world-centered representation based on the ground.

3. Camera-ground Model

Let the normal of the ground plane be expressed by $\mathbf{n_g}$ in the viewer-centered coordinate system. $\mathbf{n_g}$ actually implies the relation between the perceiver and the ground. It is not difficult to find by introspection that we usually assume the following relation to the ground. Suppose that the camera is originally so set that the optical axis of the camera and the horizontal axis of the image plane are parallel to the ground, as humans look forward while keeping two eyes horizontal. Rotate the camera around the horizontal axis of the image plane by an angle of α (see Figure 1)

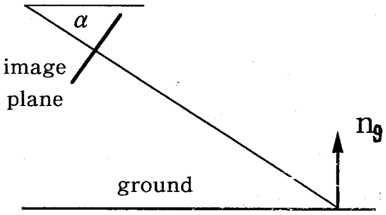


Figure 1 The camera-ground model

as humans look some feet ahead on to the road. Then the normal of the ground plane is projected to be upright onto the image under orthographic projection; i.e.,

$$\mathbf{n_g} = (0, 1, -tan\,\alpha), \qquad (1)$$

To keep $\mathbf{n_g}$ upright everywhere, which is desirable, it is necessary to assume the orthographic projection, if $\alpha \neq 0$.

4. Ground Contact Relations

Because of the planarity of rectangle and the linearity of its sides, there exist only three kinds of contact relations between the ground and a rectangle; (1) the whole rectangle contacts the ground; (2) one of the four sides contacts the ground; and (3) only one of the four corners contacts the ground. When the whole rectangle contacts the ground, the orientation of the rectangle and that of the ground are identical. All the four sides are perpendicular to the ground. When only one side contacts the ground, that side is perpendicular to the normal of the ground. If the rectangle does not stand vertically, it is interpreted to be prevented from falling by something else behihnd it. When only one of the corners contacts the ground, it is most likely that we perceive the rectangle standing vertically.

Which contact relation is perceived is largely dependant on the assumption of the perceiver's posture. The full contact relation is perceived only if the upper and lower corners are obtuse angles much greater than 90 degrees (Figure 2a) —i.e., the rectangle is remarkably slanted towards the sky — because we are not used to looking downward (see [Stevens, 1981] for an analysis of the relation between the image angle of two orthogonal space vectors and the orientation of their outer product.) The corner contact relation is perceived if one ofthe diagonal, of which the midpoint is the centroid, is vertical in image, and the upper and lower corners are acute angles (Figure 2b) — again because of the posture assumed by the perceiver. The one side contact relation is perceived if the full contact and the one corner contatc relations are not. The side that has the smaller angle to the horizontal axis is most likely perceived to contact the ground plane, because we prefer interpretations that are less slanted from the image plane (Figure 2c,2d).

We do not intend to claim the completeness of the analysis, because perception od one contact relation is not necessarily exclusive of another and the conditions are not completely quantified. Even so, however, if we manage somehow to quantifyt the conditions — e.g., for the full contact relation, the condition may be that the upper and lower angles are greater than 150 degrees — then we can completely determine the contact relations.

The above three contact relations can be respectively expressed as

$$\mathbf{PF} \cdot \mathbf{n_g} = 0, \quad \text{and} \quad \mathbf{QF} \cdot \mathbf{n_g} = 0; \tag{2}$$

$$(\mathbf{PF} \times \mathbf{QF}) \cdot \mathbf{n_g} = 0; \tag{3}$$

$$\mathbf{PF} \cdot \mathbf{n_g} = 0, \quad \text{or} \quad \mathbf{QF} \cdot \mathbf{n_g} = 0. \tag{4}$$

5. Summary

We propose in this paper the gravity regularity to be an active part in the process of interpreting quadrilaterals. Relation between the perceiver and the ground, and relation between the ground and the rectangles to be perceived, and constraints to determine rectangle orientations have also been derived. Experiments will be done in the future.

References

Barnard, S. and Pentland, A. (1983) Three dimensional shape from line drawings, **Proc. 8th Int. Joint Conf. on Artificial Intelligence**, pp. 1061-1063

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

Brady, M. and Yuille, A. (1984) An extremum principle for shape from contour, **IEEE Trans. on Pattern Analysis and Machine Intelligence**, Vol. 6, pp. 288-301

Kanade, T. and Kender, J. (1983) Mapping image properties into shape constraints,: Skewed symmetry, affine-transformable patterns, and the shape-from-texture paradigm, **Human and Machine Vision**, ed. by Beck, Hope and Resonfeld, Academic Press

Pentland, A. (1986) Introduction: From Pixels to Predicates, **From Pixels to Predicates**, ed. by Pentland, A., Ablex

Reuman, S. and Hoffman, D. (1986) Regularities of nature: The interpretation of visual motion, **From Pixels to Predicates**, ed. by Pentland, A., Ablex

Sedgwick, H. (1987) Layout2: A production system modeling visual perpspective information, **Proc. 1st Int. Conf. on Cmputer Vision**, pp. 662-666

Stevens, K. (1981) The visual interpretation of surface contours, **Artificial Intelligence 17**, pp. 47-73

Tsuji, S. Zheng, J. and Asada, M. Stereo vision of a mobile robot: World constraints for image matching and interpretation, **Proc. IEEE Int. Conf. Robotics and Automation**, pp. 1194-1199

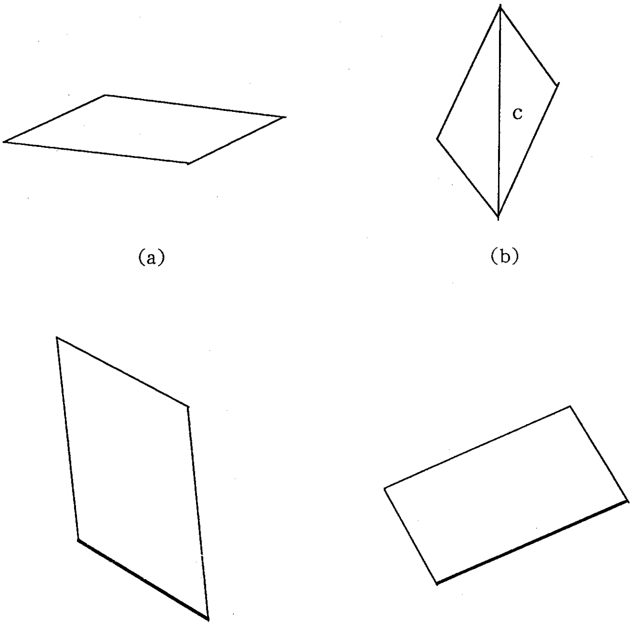
Ullman, (1979a) **The Interpretation of Visual Motion**, MIT Press

Ullman, (1979b) The Interpretation of structure from motion, **Proc. of the Royal Society, London, B203**, pp. 405-426

Xu, G. and Tsuji, S. (1987a) Inferring surfaces from boundaries, **Proc. 1st Int. Conf. on Computer Vision**, pp. 716-720

Xu, G. and Tsuji, S. (1987b) Recovering surface shape from boundary, **Proc. 10th Int. Joint Conf. on Artificial Intelligence**, pp. 731-733

Xu, G. and Tsuji, S. (1988) The visual interpretation of quadrilaterals: The Gravity Regularity, **JSIP, CV-56**



(a)

(b)

(c)

(d)

Figure 2 (a) the full contact, (b) the corner contact, and (c) and (d) the one side contact relations