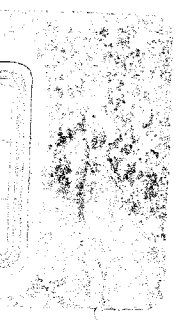
THE VISUAL INTERPRETATION OF QUADRILATERALS

GANG XU

FACULTY OF ENGINEERING SCIENCE
OSAKA UNIVERSITY

December 1988



THE VISUAL INTERPRETATION OF QUADRILATERALS

By
Gang Xu
December 1988

A DISSERTATION
SUBMITTED TO
THE FACULTY OF ENGINEERING SCIENCE
OF OSAKA UNIVERSITY
IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Contents

Preface

Computer Vision is a research field that encompasses many topics, each of which tries to tackle a particular problem. They include stereo, motion, texture, color, contour, and so on. This thesis is generally concerned with the problem of the visual interpretation of line drawings, or as sometimes people call it, the problem of shape from contour. Actually they are not exactly the same, though the two largely overlap. The problem of shape from contour is still too big to tackle in one step. Thus what we did was to choose some target which is within our reach, and by shooting it we can obtain some insights into the nature of the general problem of interpreting image contours. The target we chose is quadrilaterals, planar and curved, with which everybody gets familiar in his/her very early stages of education.

The work described in this thesis was done during my three-year Doctor's Course in Osaka University. Just before it I took the prerequisite two-year Master's Course, during which I studied stereo, also a subfield of Computer Vision. I came to Japan in 1983 after graduating from the Nanjing Institute of Technology (now Southeast University) in China. I was very fortunate to be able to work under the guidance of Professor Saburo Tsuji. I was much indebted to him for his intellectual support. He provided me with inspiring advice and timely encouragement, trusted me with the freedom to do my own thing, and talked with me not only about computer vision but also about society and people.

I thank Professor Seiji Inokuchi and Professor Tadahiro Kitahashi for serving on my reading committee. The thesis has benefited from their comments and suggestions. Professor Inokuchi also helped me receive a scholarship from the Okumura Foundation, which supported me a great deal.

Professor Takeshi Kasai, since my first day in the university, has always given

me delightful time. I really enjoyed talking with him and listening to his historical stories, not only Japanese ones but also Chinese ones.

I am grateful to Dr. Masahiko Yachida, Dr. Minoru Asada, Dr. Norihiro Abe and Dr. Masakazu Imai for their constructive suggestions and generous coorperations.

I was really fortunate to have Hayase-san (Toshio Hayase) in the laboratory, who has been always there whenever I needed him. Many others contributed to the working environment of the laboratory. At the risk of forgetting some, they include Hengli Guo, Mayumi Akimoto, Nobuki Kajihara, Hiromi Tanaka, Weon Geun Oh, Seiji Yamada, Jiangyu Zheng, Qian Chen, Ichiro Kimura, Yasuaki Okamoto, Shunji Sako, Hideki Kondo, Kazuhisa Nishimura and Yasuo Hirouchi. My thanks also go to Hiroshi Kaneko and Takeshi Shakunaga for their kind hospitality during my 5-week stay at the Musashino Telecommunications Research Institute of NTT in the summer of 1984.

Last but not least, I would like to thank the Chinese people, who bore the expense of my five-and-half-year stay in such a costly country.

Chapter 1 Introduction

1.1 Philosophy of vision research

Human brain, including the perceptual systems, is probably the last hardest nut for scientists to crack. It is so complex that one cannot really understand how it works by just cutting it into pieces, because it is of little value to observe the "dead" neural networks. Neurophysiologists actually do cut the brain into pieces to investigate how the neurons connect to each other, but this does not directly tell us what and how the brain is doing. The situation is analogous to the relation between physics and chemistry; one cannot understand all the chemical properties of a substance by just studying its physical properties. Besides the physiological approach, we have the psychological approach and the computational approach to brain research. Let us constrain our attention to vision. There is a division of labor for the three approaches. The computational approach tries to reveal what is computed and how it is derived; the psychophysical approach studys what specific algorithms are employed to calculate the solution; and the neurophysiological approach analyzes how the algorithms are implemented by the neural structures.

Similarly, in computer vision, we also have three levels of research: the computational theory, algorithm and implementation [Marr, 1982]. At the level of computational theory, human vision and computer vision are considered to do the same thing. At the level of algorithm, there usually are many algorithms to solve any particular equation, and human vision and computer vision do not necessarily employ the same algorithms. At the level of implementation, of course, the hardware of human vision and that of computer vision are completely different.

The three levels — computational theory, algorithm and implementation — can be completely independent [Marr, 1982], or largely interdependent. Although addition can be accomplished equally by both an abacus and a calculator, a von

Neumann computer and a neural computer do fairly different things, with one able to do one thing and the other able to do another.

The work described in this thesis is mainly concerned with the computational end, i.e., what is involved in the derivation of information about the 3-dimensional world from the 2-dimensional figures, although we also present algorithms to solve the specified problems.

That we take the computational approach to vision research does not mean we see vison as a pure *computation* problem. On the contrary, we see vision as a process of *inference*.

The projection of our 3-D world onto a 2-D image plane is a many-to-one mapping. But the inverse — from a 2-D image to a 3-D scene is a one-to-many mapping; there are an infinite number of possible states of the 3-D world that are consistent with any 2-D image. Selecting one from the infinite number of possible interpretations is exactly what our visual perception does. This fact simply means that vision is a process of *inference*, one of inferring the most reasonable state of the world that is consistent with the given 2-D projection. It is clear, then, that additional information is required in this process to draw the conclusion. This additional information is the knowledge of the world, of the world's stable structural regularities. It is this knowledge that is the secrets of human vision and human perception at large. It fills in the blank inherent in the mapping from 2-dimensionality to 3-dimensionality. Only by understanding this knowledge, can we really understand human vision and can we further develop any computer vision systems. Our task is thus to investigate what knowledge is employed in human vision, and to put it into an artificial vision system.

In this sense, a vision system is a knowledge system. It inherits most of the properties of any knowledge system. One needs to know which piece of knowledge to apply at a particular time. One possible choice is the generate-and-test

paradigm. Thus there need to be not only bottom-up processes but also top-down ones [Winston, 1984]. Once a piece of knowledge is triggered, the attached module computes consequent solutions. In this sense vision is also a computation problem.

The knowledge of the world can be classified into two categories: inherited knowledge and learnt knowledge. The two categories of knowledge are compared with each other as follows: the former has been embedded into human visual system through millions of years of evolution, and the latter through individual's learning experiences (of course, evolution is also a process of learning;) the former mainly serves to perceive shape and the latter mainly serves to perceive semantics, because semantics is learnt after birth; the former is common among individuals of a species, and the latter is relatively personal; the former is general and the latter relatively specific; the former is unconscious and the latter conscious.

This thesis is mainly concerned with the perception of shape, thus the knowledge we are interested in is the inherited knowledge. Since it is unconscious, we cannot know it by introspection. The way we know it is through logical reasoning. Examples of it are rigidity in motion [Ullman, 1979a,b; Reuman & Hoffman, 1986], continuity in stereo [Marr & Poggio, 1979; Grimson, 1981] and transversality in part decomposition [Hoffman, 1983; Hoffman & Richards, 1986].

On the other hand, the human visual perception is not simply a copy of external world; it has its own internal structure, which functions in its own right. One indication of such an internal structure is the perceptual system's preference of *prägnanz*, or the preference of simple, regular forms over complex, irregular ones. It is unfair to contribute this property to regularities of external world. Circles are preferred over ellipses [Barrow & Tenenbaum, 1981; Brady & Yuille, 1984] not because we regularly see circles in our environment (even though we see them more frequently,) but mainly because the internal structure of the perceptual sys-

tem appreciates the simplicity or figural goodness of the circle interpretation. If we call the knowledge of our external world the *natural regularities*, then we should call the knowledge of our internal perceptual system the *subjective regularities*. (In fact, the preference of simplicity is not the privilege of perception; all the phases of cognition show this tendency [Kanizsa, 1979, p.238].)

1.2 What is a line drawing?

Computer Vision is a research field that encompasses many topics, each of which tries to tackle a particular problem. They include stereo, motion, texture, color, contour, recognition, *etc.* This paper is generally concerned with the problem of interpreting line drawings, or as sometimes people call it, the problem of shape from contour. Actually they are not exactly the same, though the two largely overlap. Line drawings have surprisingly great descriptive power. Barrow and Tenenbaum (1981) present an informal experiment to show that different shading gredients have little effect on the perceived shape of a surface defined by a given outline. Biederman (1985) also show that simple line drawings depicting complete objects can be identified almost as quickly as full-colored slides of the same objects. This experiment reinforces his premise that the earliest access to a mental representation of an object can be modeled as a matching of a line drawing representation of a few components.

Let us first catch what the term *line drawing* means. Line drawing can be defined in two different ways. One way is to define it from image; i.e., a line drawing is a representation of the edges, or intensity discontinuities, in an image, which result from depth discontinuities, orientation discontinuities, reflectance discontinuities and illumination discontinuities. The other way is to define it as drawn by hands; i.e., a line drawing is a two-dimensional picture composed of only line segments, invented by humans as a means of representing, describing and communicating three-dimensional shape of objects.

The two kinds of line drawing are the same in the sense that they both represent significant changes in object shape, but they have some major differences. First, the hand-drawn figures are first filtered through human brain and thus include only semantically meaningful contours. But the edge images, extracted from real images by computer, include not only significant features representing shape changes, but also reflectance changes and illumination changes, and more seriously, they may also include many noise edges, which are inevitable if not processed very carefully. Secondly, the contours in a hand-drawn figure are not necessarily positionally accurate, but are topologically well-defined. On the other hand, the contours in an edge image are the honest projections of significant changes in the scene, but because of noise edges, the topological structures of the edges are not clear enough to start interpretation straightaway.

The line drawings we concern ourselves with in this thesis are those which include only topologically well-defined, semantically meaningful, but not necessarily positionally accurate (of course it is better to be accurate) contours. They can be generated either from hand-drawn figures or from edge images. In the former case, what we need to do is to straighten the "straight" lines and to smooth out the "smooth" curves so that a computer can recognize them. In the latter case, we need to discard abundant meaningless noise edges and to supplement edges that are vital but absent due to the properties of edge extraction filters.

1.3 Review and critics of previous work

There have been generally two major approaches towards interpreting line drawings. The first is the line labelling approach, and the second is the "shape-from-contour" approach, to which the work described in this thesis belongs.

The first approach was initiated by Huffman [Huffman, 1971]. He restricts the scene to be a mini-world of trihedral polyhedra. A taximony of line labels describing the edges — convex, concave and occluding — is proposed. Using this taximony, each line junction (or vertex) is given a single catagorization defined in the junction dictionary, if the line drawing is a legal one. Unfortunately, the existence of such a globally consistent line labeling is only a necessary but not sufficient condition for the correctness of the line drawing. As the same junction dictionary was also independently discovered by Clowes [Clowes, 1971], the scheme is called the *Huffman-Clowes labeling scheme*. The work was followed by a series of papers that try to extend it. Waltz (1975) includes shadow and crack edges in the labeling set. Turner (1974) and Chakravarty (1979) entend the world to include curved objects. Sugihara (1979) develops a junction dictionary to deal with range data of objects of planar and curved faces. Mackworth (1973) proposes *gradient space* that is modified from the idea of *dual space* in [Huffman, 1971] and uses it to derive necessary conditions for the correctness of a line drawing. Sugihara (1986) proposes the use of linear algebra to derive necessary and sufficient conditions for a line drawing to represent a legal scene. His approach reduces the correctness problem to one of checking the existence of a solution to a set of linear equations and inequalities. Kanade (1980) develops a junction dictionary to deal with the *Origami* world, or paper-made objects.

While the line labeling approach does not produce any quantitative results, the shape-from-contour approach tries to recover surface shape from line drawing by employing additional constraints. An early attempt was the *general viewpoint*

assumption (see [Binford, 1981].) By assuming it, the most stable interpretation — it means that a slight change of viewpoint does not drastically change the image configuration and its interpretation — is selected from among the possible. We will discuss it later in detail.

Besides this, a number of other papers have also been published that try to answer the "shape from contour" question. [Koenderink & van Doorn, 1982; Koenderink, 1984] provide a mathematical (differential-topological) analysis on the relations between extremal boundaries and the underlying surfaces. Others are [Barrow & Tenenbaum, 1981], [Kanade, 1981], [Barnard, 1983] and [Brady & Yuille, 1984], which generally deal only with closed contours and assume that the image contours are, globally or locally, projections of planar space curves. Ellipses are interpreted as circles by additional assumptions of: *uniformity of curvature* [Barrow & Tenenbaum, 1981], *maximum entropy* [Barnard, 1983] and *maximum compactness* [Brady & Yuille, 1984], and quadrilaterals or parallelograms are interpreted as rectangles by additional assumptions of *maximum symmetry* [Kanade, 1981; Barnard, 1983]. Reviewing these papers, it is not difficult to find that the approaches are based on the psycological facts that ellipses and quadrilaterals (including parallellograms) are perceived as circles and rectangles, respectively. The approaches differ only in how the facts are accounted for and in what specific criteria they choose to achieve the predetermined aims. The psycology of line drawing perception was first studied by the Gestalt school [Kanizsa, 1979]. They propose *prägnanz*, or figural goodness, to be the criterion on which the human perception is based. The circle and rectangle interpretations are preferred because they are the most beautiful interpretations among the possible. Unfortunately, however, the Gestalt psycologists could not give a theoretical account of the term *prägnanz* from the standpoint of information processing, and figural goodness remains to be judged mainly by human eyes (note that a recent progress

is an attempt to characterize pragnanz by transformational invariance [Palmer, 1983].) The lack of a definition of *prägnanz* leaves room for proposal of various specific criteria. The *uniformity of curvature*, *entropy*, *compactness* and *symmetry* criteria are developed in, and thus well suited to the specific cases, but they are not surely universal and do not necessarily apply to other cases.

Under the circumstances, two general paradigms are available. The first is the criterion-based paradigm, in which a universal criterion is developed and interpretations are selected by maximizing or minimizing that criterion, as done in [Barrow & Tenenbaum, 1981], [Barnard, 1983] and [Brady & Yuille, 1984]. The second is the regularity-based paradigm, in which specific figural configurations that have definite interpretations are searched for and interpreted. Two examples are [Stevens, 1981], in which parallel (curved) contours are interpreted as lines of curvature on a cylindrical surface, and [Barnard & Pentland, 1983], in which elliptic arcs are directly interpreted as circular arcs. The causal relation between the interpreted and the interpretation is referred to as regularity. The task is thus to discover regularities, or causal relations, and then to apply them to specific interpretation processes. We consider that the regularity-based paradigm is more advantageous because (1) a line drawing is generally only qualitative, not quantitative, especially when it is hand-drawn; and (2) while interpretations are qualitatively stable, they are not always quantitatively invariant.

1.4 Overview of this dissertation

The problem of line drawing interpretation is too big to tackle in one step. Thus what we did is to choose some target which is within our reach, and by shooting it we can obtain some insights into the nature of the general problem of interpreting image contours. The target we chose is quadrilaterals, planar and curved, with which everybody gets familiar in his/her very early stages of education.

This thesis is composed of 7 chapters. In Chapter 2 we first explain basic terms and concepts, then discuss what the *general viewpoint assumption* means in the specific case of inferring 3-D continuity and discontinuity of a curve, and finally apply it to derive two important specific constraints on interpreting image vertical and horizontal lines.

In Chapter 3, we propose the rectangularity regularity to be the prime constraint in the visual interpretation of quadrilaterals and discuss the nature of the regularity. Interpreting quadrilaterals as rectangles under both orthographic and perspective projections is accounted for in a unified and consistent way. The "image center" and focal length are determined together with rectangle orientation.

In Chapter 4, the rectangularity regularity is adapted to apply to "curved quadrilaterals". Applying the line of curvature (LOC) regularity, a 4-segment closed image contour is interpreted as four 3-D lines of curvature of the surface the image contour depicts. We further propose a segmentation rule, an algorithm for constructing a net of lines of curvature inside the boundary, and a method for estimating surface orientations at the net intersections. Some experimental results are also presented.

In Chapter 5, we analyze the role that gravity plays in visual perception and in interpreting quadrilaterals especially. 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. We first analyze the relation among the perceiver (camera), the ground and the rectangles supported by the ground, and then derive constraints to determine the rectangle orientation.

In Chapter 6, we report the conditions for interpreting quadrilaterals as faces of a rectangular polyhedron. There are two levels of interpretation: the geometrical level and the perceptual level. The differences between interpretations at the two levels are also examined.

In Chapter 7, we present a brief review of the thesis, elaborate on some points, and make some concluding remarks.

Chapter 2 The General Viewpoint Assumption

2.1 Projection and the GVP assumption

We start with the introduction of the camera model. As shown in Fig. 2.1, F is the focus, which has the coordinates $(x0, y0, -f)$. The focal length is f. The image plane is the xy-plane, and the image center O has the coordinates $(x0, y0)$. We intentionally make both the "image center" and the focal length unfixed, because they are to be determined in the process of interpretation. While the camera system is a real one if the image is, it is a subjective one if the image is a hand-drawn figure.

An arbitrary space point (x, y, z) is projected onto the image plane as

$$\begin{cases} (\, x0 + (x - x0)\frac{f}{(f+z)},\ y0 + (y - y0)\frac{f}{(f+z)}\,), & \text{if } f \neq \infty; \\ (\, x,\ y\,) & \text{if } f = \infty. \end{cases}$$

If f is finite, then the projection is called *perspective projection*; if f approaches infinity, then it is called *orthographic projection*. We do not have real cameras that have a focal length approaching infinity, but people usually assume for simplicity orthographic projection as an approximation if object dimensions are small compared to the distance to the objects. In the case of hand-drawn figures, however, some are produced by assuming orthographic projection, and evidently different from images taken by a real camera. Thus they can be interpreted only by assuming the "impossible" orthographic projection again.

The *general viewpoint* assumption is the most essential assumption in interpreting line drawings. Without it one cannot move even a tiny step. A straight line in image is not necessarily the image of a straight line in space; it can be a planar curve when the viewing direction is on that plane. It is also not necessarily continuous; it may be broken segments. It is interpreted as a straight line in space only when the general viewpoint is assumed. The GVP assumption has

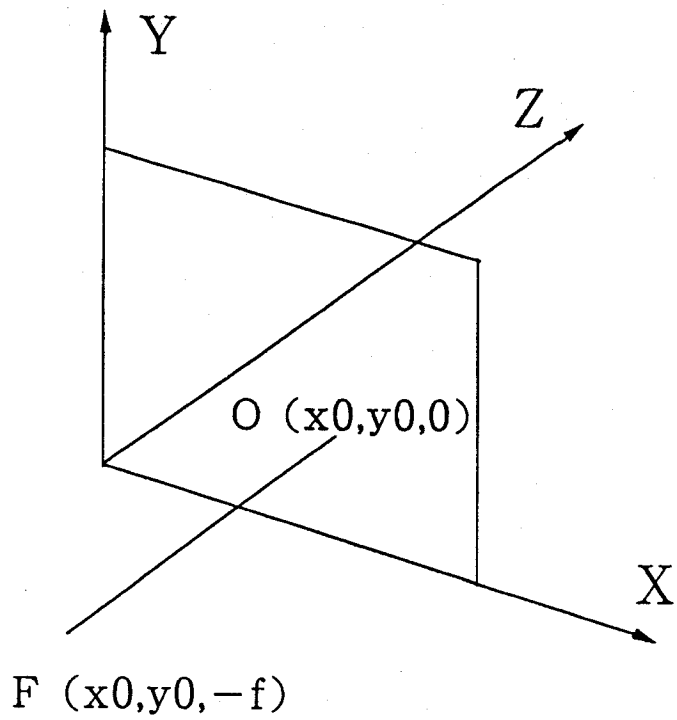


Fig. 2.1 The camera coordinate system

been applied to many specific cases to derive specific constraints. Binford (1981) successfully presents a series of costraints which he argues should be invoked in absence of other evidence. Stevens (1981) also applies the GVP assumption to reason that parallel image curves should be interpreted as lines of greatest curvature on a cylindrical surface.

Conventionally, the *general viewpoint* is literally explained as such a position that perturbation of it does not qualitatively change the line-drawing configuration. However, I would like to give another definition; i.e., the *general viewpoint* is such a camera position that a small change of it does not cause any degradation, or, a small change of it does not make a more structured configuration less structured (e.g., a straight image line does not become a curved line, or, a continuous line does not become broken, etc.) Although the terms *"unstructured"* and *"structured"* still lack precision from the standpoint of information-processing, this definition explicitly reveal in what direction the perturbation of camera from a non-general position might change the image configuration, and in terms of "structuredness" this definition opens a possibility for the GVP assumption to be assimilated into the general perception law *prägnanz.* (Detailed analyses of *prägnanz* can be found in [Witkin & Tenenbaum, 1983; Witkin & Tenenbaum, 1986; Kanizsa, 1979; Palmer, 1983], and a detailed analysis of "degeneracy" can be found in [Kender & Freudenstein, 1987].) In the following two sections we will apply the GVP assumption to analyze the relations between image continuity and discontinuity and corresponding space continuity and discontinuity, and to account for the well-known fact that vertical and horizontal image lines are perceived as vertical and horizontal space lines, respectively.

Incidentally, we clarify the meanings of the terms *regularity*, *assumption* and *constraint*, which are often used, and sometimes confusingly. They essentially refer to the same thing, but differ in nuance. *regularity* is used in the objective

sense; *assumption* is used in the subjective sense; and *constraint* is used in the functional sense, once a regularity is assumed. We inherit the original uses if they are proposed by the others, e.g., the *general viewpoint* assumption [Binford, 1981], the continuity constraint [Marr & Poggio, 1979; Grimson, 1981] and the transversality regularity [Hoffman & Richards, 1986].

2.2 The GVP assumption for continuity and discontinuity

The terms we will use through this section, *1st-order, 2nd-order* and *3rd-order* continuity and discontinuity, mean continuity and discontinuity in position, tangent and curvature, respectively. A higher order continuity implicitly implies continuity of lower orders. For simplicity, when we say an nth-order continuity we mean that n is the highest order of continuity and at the same time the nth-order continuity is a $(n+1)$th-order discontinuity. For example, the intersection of two straight lines is a 1st-order continuity and at the same time a 2nd-order discontinuity.

It is trivial that 3-dimensional continuities in position, tangent and curvature are projected onto the image plane of a camera located at any position as at least 2-dimensional continuities in position, tangent and curvature, respectively. But for some special viewpoints, it is possible that a 3-D nth-order continuity is projected as a 2-D $(n+1)$th-order continuity. For instance, a 3-dimensional (1st-order) continuous curve is always projected at least as a 2-dimensional (1st-order) continuous curve, and at some special viewpoints two curves not intersecting in space, i.e. a 1st-order discontinuity, may be projected as coincident in image, i.e. a 1st-order continuity.

A general viewpoint, in the specific case of continuity and discontinuity, is such a viewpoint that a small change of it does not reduce order of continuity. Table 2.1 shows all possible transitions of order (from 0 to 3) of continuity and discontinuity in the process of projection, where 0th-order continuity is a replace-

ment of 1st-order discontinuity. And examples are shown in Fig. 2.2 to instantiate each transition that is caused by a special viewpoint.

While many possible order transitions of continuity are due to special viewpoints, which will not occur if the general viewpoint is assumed, some are not caused by special viewpoints (within a certain range,) but by occlusion. Two examples are the well-known *T-juncion* and *cylindrical junction* as shown in Fig. 2.3, to which [Binford, 1981] pays much attention. We here reexamine them under our new framework. A T-junction in image (Fig. 2.3a) is a 1st-order continuity plus a 3rd-order continuity (orders higher than 3 are neglected because human perceptual system is not sensitive to it and does not make use of it.) Three cases are possible: (1) the stem is nearer to the viewer, and the 2-D 1st-order continuity is due to the special camera position; (2) the stem is coincident in space; (3) the stem is further away, and only a part of it is visible. By the general viewpoint assumption, the first possibility is eliminated. The second possibility is not an order transition. The third possibility is an order transition, but is not dependent on viewpoint within a certain range. Thus the junction can be perceived as a 3-D 1st-order continuity or as a 3-D 1st-order discontinuity (but the stem is further away.)

A *cylindrical junction* in image is a 2nd-order continuity plus a 3rd-order continuity (P in Fig. 2.3b,) or simply a 3rd-order continuity (Q in Fig. 2.3b.) The straight segment (let us say it without proof) in space can projectively be any segment lying on the plane determined by the focus and the image segment. Only if the straight segment is also tangent to the curves in space, can a small change of viewpoint not cause a sudden collapse of the configuration. If, however, the segment is not an *apparent edge* but an *extremal boundary*, it moves with the viewpoint and the order transition does not disappear by a small change of viewpoint (see Chapter 4 for a practical example.)

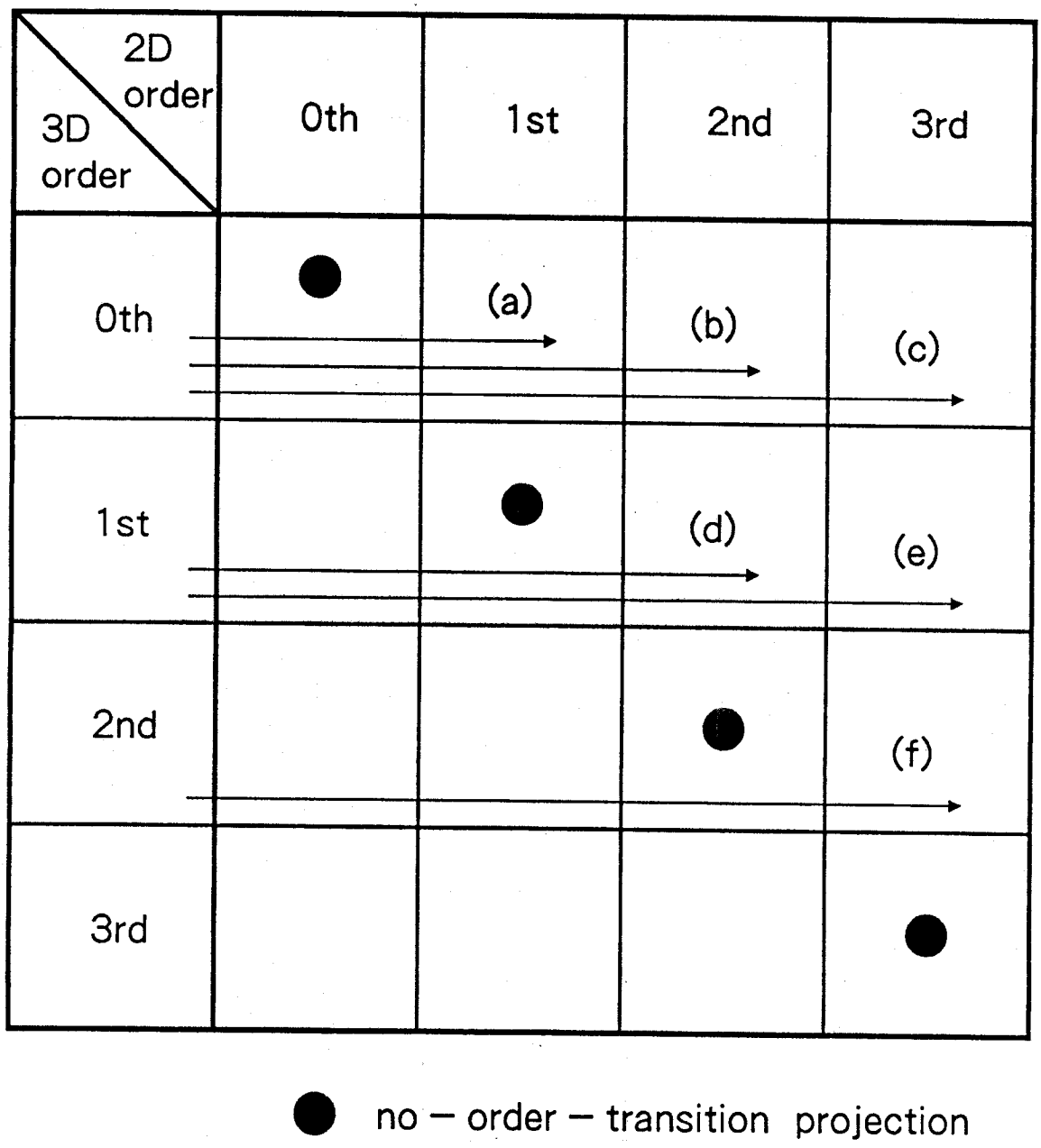


Table 2.1 Possible order transitions

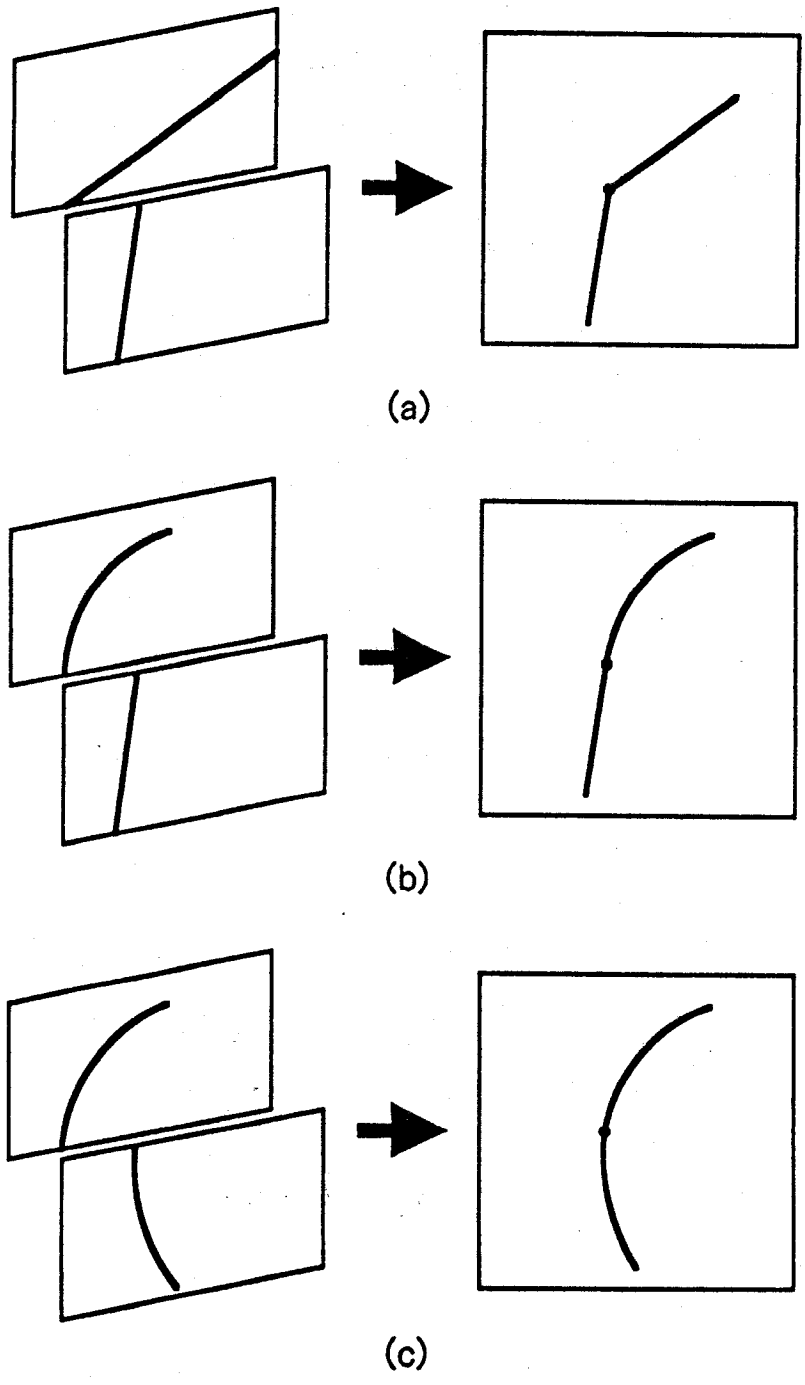
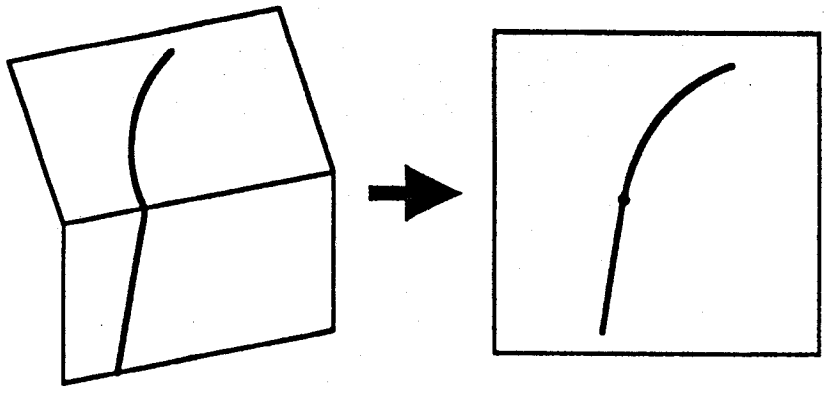
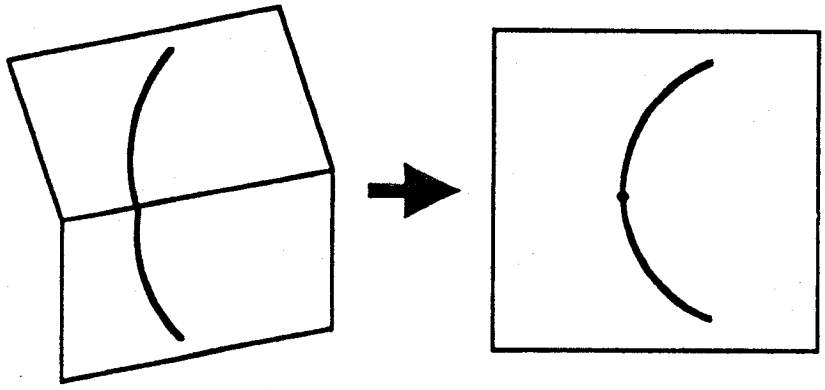


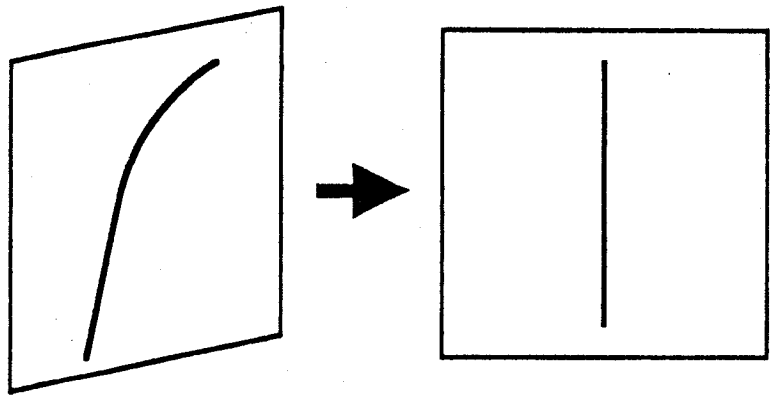
Fig. 2.2 Examples of continuity order transitions caused by special viewpoints



(d)



(e)



(f)

Fig. 2.2 (continued)

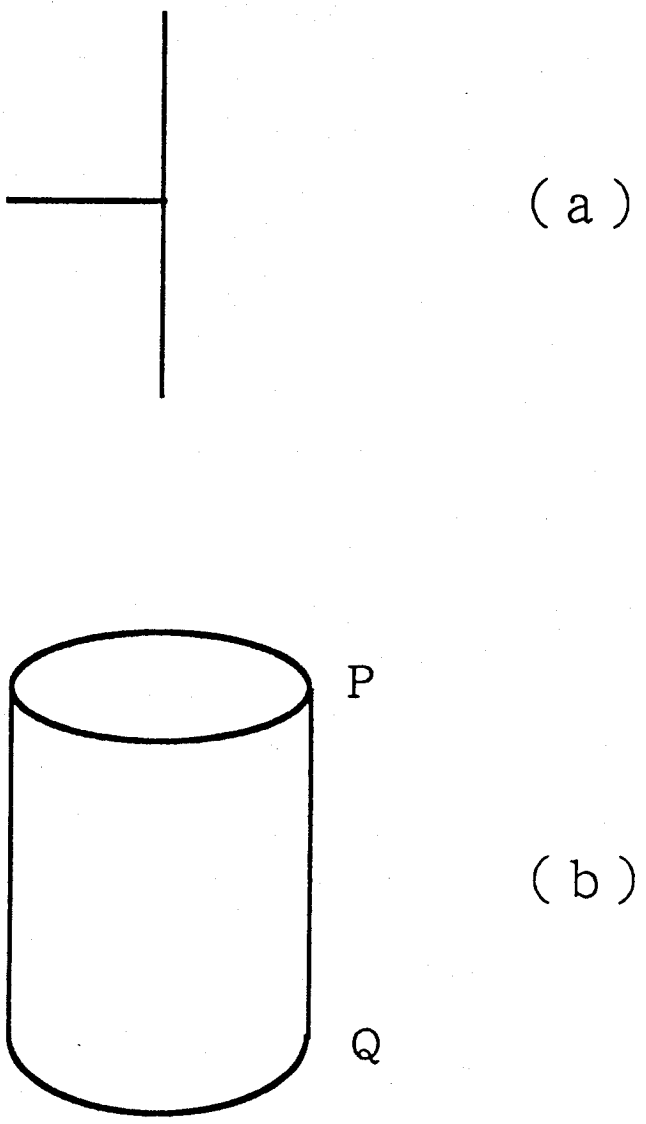


Fig. 2.3 The *T-junction* (a) and the *cylindrical-junction* (b)

In summary, in the process of projection the orders of 3-D continuity and discontinuity remain the same or increase. In the latter case, it may be dependent on viewpoint or independent of viewpoint. By assuming a general viewpoint, those order transitions due to special viewpoints can be removed from consideration. On the other hand, the number of order transitions that are independent of viewpoint is limited, and thus makes it easy to recognize them. As a result, 3-D continuities and discontinuties are definitely somehow evidenced by 2-D continuities and discontinuties. In this sense the projection is a qualitatively equivalent mapping. It is this qualitatively equivalent mapping that makes it possible to infer 3-D features from their 2-D images.

2.3 The GVP assumption for interpreting vertical and horizontal lines

It is broadly known that (in absence of other evidence) vertical image lines are perceived as vertical space lines and horizontal image lines as horizontal space lines. The reasoning is divided into two steps. Applying the GVP assumption, the first step explains why straight image lines are perceived as straight space lines and the projection can be either perspective or orthographic, and the second step reasons that vertical image lines are perceived as vertical space lines, and horizontal image lines as horizontal and image-plane-parallel space lines, and the projection must be orthographic.

Straightness means that curvature is zero. Curvature is everywhere zero along a straight line. If the back image of a straight image line has any 1st-order discontinuities, any 2nd-order discontinuities, or the curvature is not everywhere zero, a slight change of viewpoint will make the straight line unstraight. Thus by assuming general viewpoint, all these possibilities are eliminated, and the only possibility is that the back image is also straight. The property of zero curvature is irrelevant to the foreshortening. Thus the straightness interpretation is independent of the focal length; i.e., the projection can be either perspective or

orthographic.

The concepts of *verticality* and *horizontality* originate from *gravity* (description of incorporating gravity into perception as a constraint is to be detailed in Chapter 5.) 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. To keep stability, the horizontal axis of the image plane is favorably parallel to the ground, as humans look forward while keeping two eyes horizontal. The optical axis of the camera points obliquely to the ground, as humans look some feet ahead on to the road. For a space vertical line at an arbitrary location to be projected as an image vertical line, it is necessary that the projection be orthographic. Otherwise we need the condition that the camera's optical axis is horizontal or that the space vertical line is coplanar with the camera's vertical axis.

By assuming this camera model, it is clear that any space lines lying on any vertical planes, not only the space vertical lines, are projected as vertical lines in image. But if the lines are not the vertical lines, then a small change of viewpoint makes their images no longer vertical. Only the space vertical lines, even if the viewpoint changes, are still projected as vertical.

The same line of reasoning applies to horizontal image lines. Assuming the same camera model, it is clear that any space lines that are perpendicular to the image vertical axis are projected as horizontal lines. But if the lines are not simultaneously horizontal in space, then a small move of viewpoint along the vertical direction make their images no longer horizontal. Only the horizontal and image-plane-parallel space lines are still projected as image horizontal lines, even if the viewpoint translates in any direction and rotates around the horizontal axis, but not around the optical and vertical axes (the image horizontal lines will

no longer be horizontal if one turns viewing direction away.) This restriction on viewpoint change shows that a *general viewpoint* is not always absolutely general, but sometimes only means the least probability of qualitative changes due to viewpoint changes.

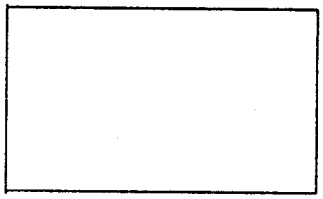
Chapter 3 Interpreting Quadrilaterals: The Rectangularity Regularity

In this chapter we propose the rectangularity regularity to be the prime constraint in the visual interpretation of quadrilaterals and discuss the nature of the regularity. Interpreting quadrilaterals as rectangles under both orthographic and perspective projections is accounted for in a unified way. The "image center" and focal length are determined together with rectangle orientation.

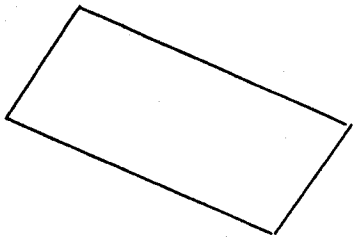
3.1 The Rectangularity Regularity

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. On the other hand, the human visual perception is not simply a copy of external regularities; it has its own internal structure, which functions in its own right. One indication of such an internal structure is the perceptual system's prefernce of *prägnanz*, or simple, regular forms over complex, irregular ones. It is unfair to contribute this property to regularities of external world. Circles are preferred over ellipses [Barrow & Tenenbaum, 1981; Brady & Yuille, 1984] not because we regularly see circles in our environment, but mainly because the internal structure of the perceptual system appreciates the simplicity or figural goodness of the circle interpretation. This kind of regularities — if one would also like to call them regularities — is subjective regularities, in contrast to external natural regularities.

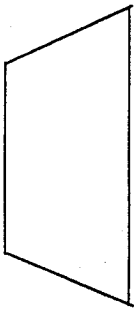
The rectangularity regularity is such a subjective regularity resulting from the internal structure of the perceptual system. All quadrilaterals can be queued, according to the degree of regularity, as: rectangle, parallelogram, trapezium and generic quadrilateral (Fig. 3.1). It is generally difficult to define degree of regularity universally, but here it is intuitive and we do not try to give a theoretical definition [Palmer, 1983]. It is observed that a single quadrilateral tends to be



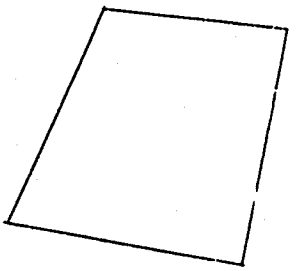
a rectangle



a parallelogram



a trapezium



a generic quadrilateral

Fig. 3.1 Quadrilterals are queued by the degree of regularity.

always perceived as a rectangle. To put it another way, a quadrilateral in image, whatever its degree of regularity is, tends to be interpreted as a rectangle in space, the most regular interpretation among the possible, and the 2-D irregularity is thought of as being caused by the projection. It is from this observation that the rectangularity regularity is generalized.

The rectangularity regularity has long been recognized as being of importance. Barnard (1983) computes the planar orientation of a rectangle from its image, a quadrilateral, under perspective projection. Kanatani (1986) discusses how to compute spatial orientations of the faces of a rectangular trihedral polyhedron. Xu and Tsuji (1987a,b) extends this regularity to curved surfaces and proposes the LOC (line of curvature) regularity to recover shape of curved surfaces. In this paper we give a unified and detailed account of interperting quadrilaterals in image as rectangles in space under both orthographic and perspective projections.

3.2 Deriving focal point and rectangle orientation

In this section we discuss how and to what extent the 3-D orientation of a quadrilateral is determined by incorporating the rectangularity regularity. The coordinate system we assume, as shown in Fig. 2.1, is different from those we usually use, in that the coordinate origin is located on the image plane and in that the z-axis is independent of the focal point, which has the coordinates $(x0, y0, -f)$. Both the "image center" $(x0, y0)$ and the focal length f are then to be determined in the process of interpretation. While the camera system is an objective one if the image is a real one, it is a subjective one if the image is a hand-drawn figure.

What is known is a quadrilateral in the image, and what is to be known is the orientation of the rectangle that projects that quadrilateral, and the location of the focal point, but not the distance or size of the rectangle. Let the four corners of the quadrilateral be A, B, C and D, as shown in Fig. 3.2. Extending the segments AB and CD, we have the intersection P $(x1, y1)$. Since AB and CD

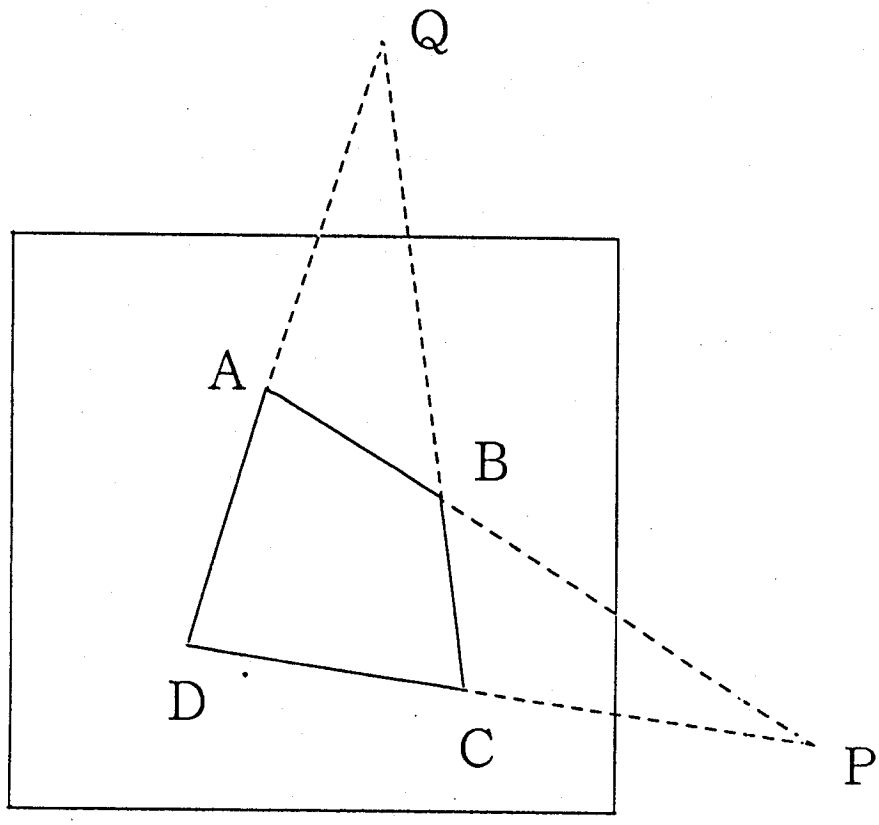


Fig. 3.2 An image quadrilateral ABCD. Extending the laterals, we have two intersection points P and Q.

are the images of two parallels, P is the vanishing point of the parallels. Similarly, extending the segments BC and DA, we have the intersection Q $(x2, y2)$, which is also a vanishing point, of the other pair of parallels. P and/or Q approach infinity if the corresponding image segments are parallel.

A straightforward demonstration of the concept of vanishing point is that the line connecting the focal point and a vanishing point is parallel to the space parallel lines that give rise to that vanishing point. As shown in Fig. 3.3, PF is parallel to A'B' and C'D', and QF is parallel to B'C' and D'A'. Since A'B'C'D' is a rectangle, PF is perpendicular to QF. The plane determined by PF and QF is parallel to the plane on which the rectangle lies, and thus the orientations of the two planes are identical. The segments PF and QF can be expressed in vector form as $(x1 - x0, y1 - y0, f)$ and $(x2 - x0, y2 - y0, f)$, respectively. From the perpendicularity, the inner product of the two vectors is zero. By this equation F can be determined or at least constrained. Once F is determined, the plane normal (let it be called n) can also be determined as the outer product of PF and QF. In the following, we discuss three cases of PF and QF: (1) neither P nor Q approaches infinity; (2) either P or Q approaches infinity; and (3) both P and Q approach infinity.

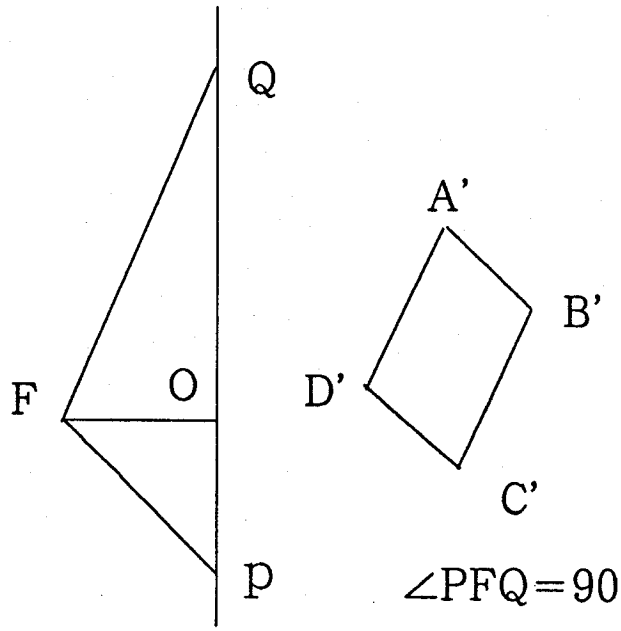


Fig. 3.3 PF is parallel to A'B' and C'D', and QF is parallel to B'C' and D'A'. Since A'B'C'D' is a rectangle, PF is perpendicular to QF.

(case1)

If both P and Q do not appproach infinity, then the quadrilateral is a generic one. From the perpendicularity of PF and QF we have

$$(x1 - x0)(x2 - x0) + (y1 - y0)(y2 - y0) + f^2 = 0. \tag{3.1}$$

Clearly, $f\,(f > 0)$ can be determined as

$$f = \sqrt{-(x1 - x0)(x2 - x0) - (y1 - y0)(y2 - y0)}. \tag{3.2}$$

This equation describes a hemisphere with a diameter PQ, on which F is constrained to lie, as shown in Fig. 3.4. For f to have a solution, the point O $(x0, y0, 0)$ must satisfy the following inequality:

$$(x1 - x0)(x2 - x0) + (y1 - y0)(y2 - y0) < 0.$$

This means that O must lie inside the circle with PQ on the image plane. When Q reaches the midpoint of PQ, all of PO, QO and FQ become radii of the hemisphere, and f has the maximal value, half the length of PQ. To determine F completely, however, we have to first know where the attention is oriented; i.e., where O is. If there are three or more quadrilaterals in a real image, then P can be determined as the intersection point of the hemispheres corresponding to each quadrilaterals (see Section 4 for a special case.) A prerequisite to this solution is that the hemispheres do have a common point; i.e., the image is a real one. However, as in hand-drawn figures, quadrilaterals are usually produced by individual attentions. Consequently, they should be, and can only be, perceived separately. One reasonable choice for each quadrilateral is the intersection of the two diagonals, which is the centroid of the corresponding rectangle in space, if the centroid is within the circle with a diameter PQ. As mentioned above, the maximal value for f is half the length of PQ. Trially, the closer to a parallelogram

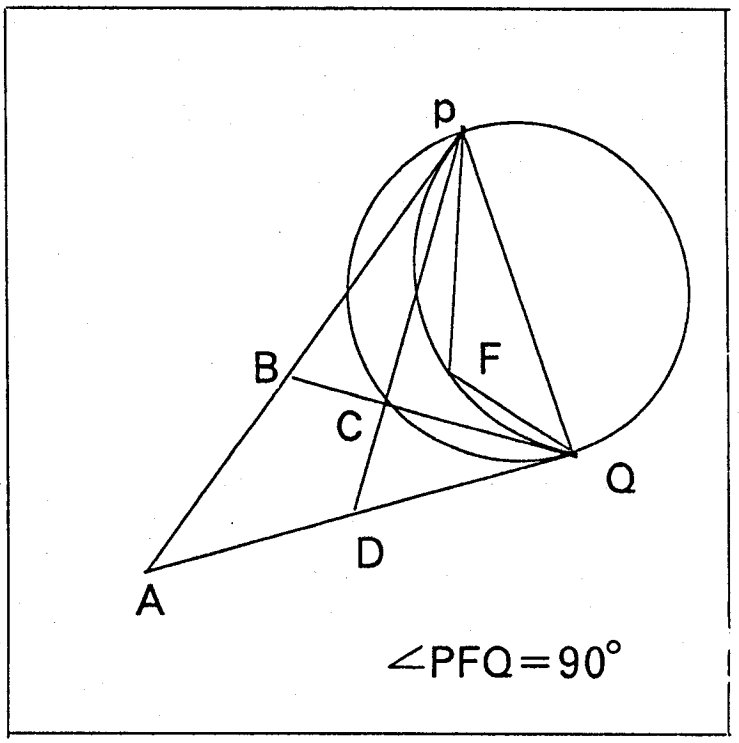


Fig. 3.4 F lies on the hemisphere with a diameter PQ.

the quadrilateral is, the greater f is. On the other hand, humans prefer long focal lengths in the perception of figures. This answers why a parallelogram is more often used to represent a rectangle, and why a parallelogram is more easily perceived by our eyes as a rectangle.

(case2)

If one of P and Q approaches infinity, then the quadrilateral is a trapezium. Without loss of generality, suppose that P approaches infinity, while Q does not. PO $(x1-x0, y1-y0)$ can be expressed by (a, b) as P approaches infinity. Adding the z-component, PF can be expressed by (a, b, c), where a, b and c are all constants. c approaches zero if f does not approach infinity, and c may still be zero even if f does approach infinity. Can f approach infinity? Suppose that it does. Then clearly BC is parallel to DA, and their intersection Q also approaches infinity. This leads to a contradiction. Thus f cannot approach infinity and c equals zero.

From the perpendicularity of PF and QF, we have

$$a(x2-x0) + b(y2-y0) = 0. \tag{3.3}$$

This equation constrains OQ to be perpendicular to AB and CD; i.e., O must lie on the line perpendicular to AB and CD drawn from O. Let us call the line IQ (Fig. 3.5). Since f is free, F is constrained to lie on the plane that is projected onto the image as IQ. When O is located at Q, the orientation is completely determined by the orientations of AB and CD as $(b, -a, 0)$.

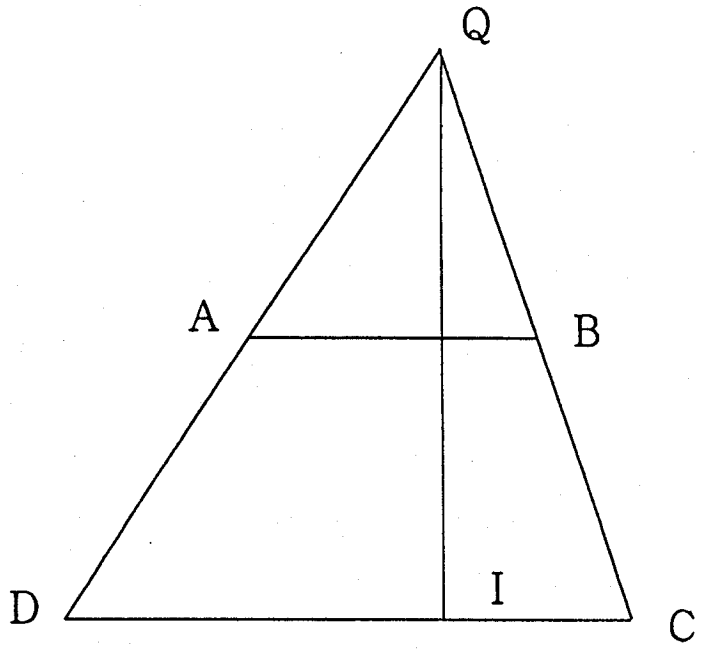


Fig. 3.5 F lies on the plane that is projected onto the image as IQ.

(case3)

If both P and Q approach infinity, then the quadrilateral is a parallelogram or a rectangle. In this case, PO is parallel to AB and CD, and QO is parallel to BC and DA. O is actually not constrained. Suppose that the orientations of PO and QO are expressed by (a, b) and (a', b'), respectively. Adding the z-components, PF and QF are expressed by (a, b, c) and (a', b', c'), respectively. a, b, c, a', b' and c' are all constants. Both c and c' approach zero if f does not approach infinity, and c and c' may still be zero even if f does appoach infinity. From the perpendicularity of PF and QF, we have

$$aa' + bb' + cc' = 0. \tag{3.4}$$

A special case is that ABCD is a rectangle; i.e.,

$$aa' + bb' = 0. \tag{3.5}$$

Then

$$cc' = 0. \tag{3.6}$$

Three cases are possible:

$$\begin{cases} c = 0, \, c' = 0; \\ c = 0, \, c' \neq 0; \\ c \neq 0, \, c' = 0. \end{cases}$$

If either of c and c' is not zero, then f approaches infinity and the projection can only be orthographic; otherwise ABCD would not be a rectangle. If both c and c' are zero, then f is not constrained. The projection can be either orthographic or perpspective. In this case, the rectangle normal can be determined as $(0, 0, 1)$; the rectangle is parallel to the image plane.

Chapter 4 Interpreting "Curved Qadrilaterals": The LOC Regularity

In this chapter we propose the line of curvature (LOC) regularity, by which a 4-segment closed image contour is interpreted as four 3-D lines of curvature on the surface that the image contour depicts. Based on the LOC regularity, we present a segmentation rule, an algorithm for constructing a net of lines of curvature inside the boundary, and a method for estimating surface orientations at the net intersections. Some experimental results are also illustrated.

4.1 Lines of curvature and the LOC regularity

For a curved surface, how the surface curves is an important description. Euler discovered that at any point on any surface there are always a direction in which the surface curves most, and a direction in which the surface curves least (spheres and planes are exceptions, because the surface curvature is identical in every direction at every point.) The two directions are always orthogonal and are called the *principal directions*, or, the *direction of greatest curvature* and the *direction of least curvature*. The corresponding curvatures are called the *principal curvatures*. By starting from some point and always moving along the direction of the greatest curvature, we trace out a *line of greatest curvature*, and similarly, by moving in the direction of least curvature, we trace out a *line of least curvature* [Struik, 1961]. The lines of greatest curvature and the lines of least curvature form two mutually orthogonal families of curves covering the surface simply and without gaps (excluding umbilic points where the curvature is equal in every direction). They are together referred to as *lines of curvature*. On a surface of revolution lines of curvature are the meridians and the circles, and on a cylindrical surface lines of curvature are the straight rulings and the parallel cross sections.

The line of curvature (LOC) regularity describes the relation between certain image contour configurations and the line of curvature interpretations. Examples

are shown in Fig. 4.1. Each closed contour is interpreted as four lines of curvature on the surface that the contour depicts. While a detailed segmentation rule is to be given in a later stage, a brief description of the configuration is given here. The closed contour is formed by four segments, the intersections are the tangent discontinuities, or the curvature discontinuities, and every two non-intersecting segments are similar each other to a certain degree.

Another example is shown in Fig. 4.2. The parallel contours are interpreted by [Stevens, 1981 & 1986] as lines of curvature on a cylindrical surface. The other family of lines of curvature is the straight rulings, which can be recovered by connecting the corresponding points on the parallel contours.

While the LOC regularity does appeal to our intuition, how can we assess that to exploit the LOC regularity is a right choice even if it is not *a priori* guaranteed valid? The situation is basically different from that of Ullman (1979); the rigidity regularity is guaranteed valid by taking enough views of enough dots to overdetermine the solution for rigid motion, but we do not have any means of judging the validity of the LOC regularity from the image contours themselves. An argument can be made on the statistical ground. Stevens (1986) suggested that the majority of physical curves across the surfaces of manufactured objects and many biological forms are lines of curvature. This line of argument does provide the existential basis for the LOC regularity, but one still cannot guarantee its validity in specific cases. Here again we consider that the LOC regularity is a subjective one, resulting from the perceptual system's preference of regular forms over irregular ones. To put it another way, the line-of-curvature interpretation is the most regular, most structured one among the possible. If this is true, then we can further conclude that lines of curvature provide more power to constrain a surface than any other curves.

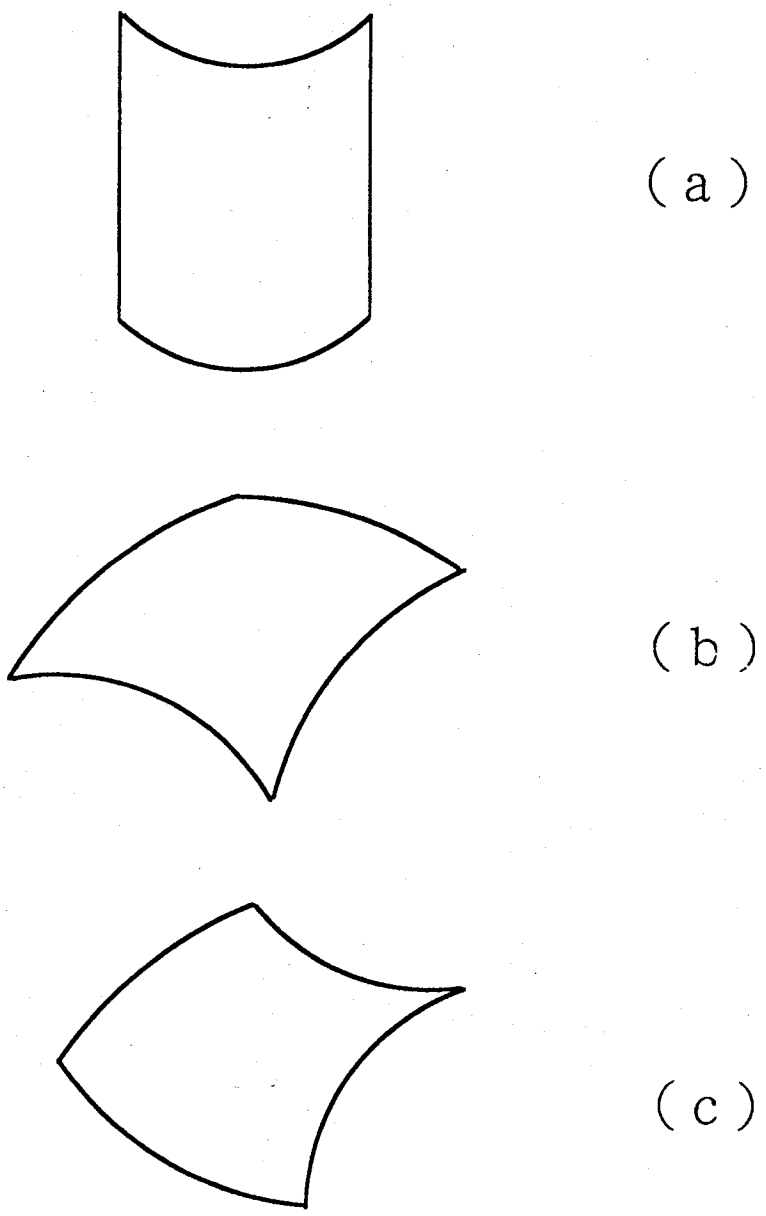


Fig. 4.1 A cylindrical (a), an elliptic (b) and a parabolic (c) surfaces with line of curvature boundaries

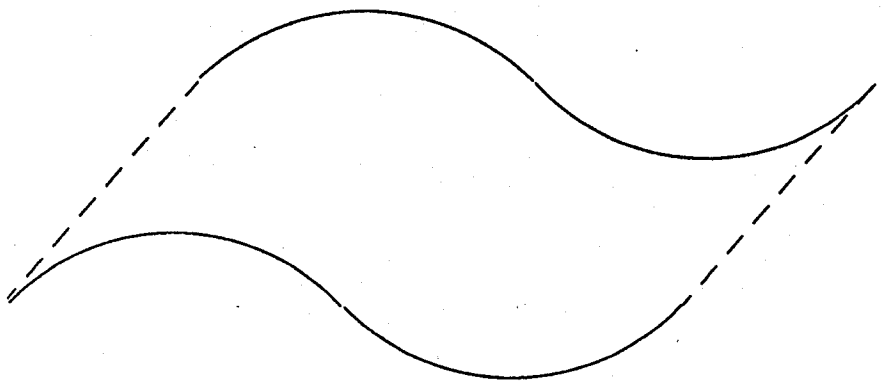


Fig. 4.2 Parallel surface contours are interpreted as lines of curvature on a cylindrical surface by Stevens (1981).

4.2 Segmentation rule

A closed boundary is segmented as lines of curvature if the following conditions are satisfied: (1) there are four tangent discontinuities or curvature discontinuities; and (2) every two non-intersecting segments defined by the four discontinuities (there are only two pairs) are similar each other.

Recall that lines of curvature intersect at a right angle in space. When projected onto image at a general viewpoint (see Chapter 2), the intersections become tangent discontinuities. The only exception is the case of *cylindrical junction*, where the intersection is on an *extremal boundary* and the line of sight divides the right angle. While the image tangent at such an intersection is continuous, the image curvature is discontinous (human eyes are insensitive to derivatives of order of 3 and more). Thus, by searching for tangent discontinuities and curvature discontinuities, we can segment a closed image boundary into continuous space segments. Fig. 4.3 is simply a copy of Fig. 2.3b.

In general, 4 lines of curvature bound a simple surface patch. Every two non-intersecting segments form a pair. Thus there are two pairs, one of lines of greatest curvature, the other of lines of least curvature. A line of greatest (or least) curvature can be thought of as obtained by moving the other line of greatest (or least) curvature along the way guided by the lines of least (or greatest) curvature. For an arbitrary smooth surface, the two segments of each pair are not parallel (both 2-dimensionally and 3-dimensionally), but are deformed smoothly, and thus similar, to some extent (in some sense the similarity is a kind of symmetry [Brady & Asada, 1984].) The similarity can be quantified by computing the correlation of the two curvature functions (of the segment points.)

While the four-segment boundaries are a general case, three-segment and five-segment boundaries can also be interpreted as lines of curvature in certain cases. For example, a conical surface may have only three lines of curvature, if the

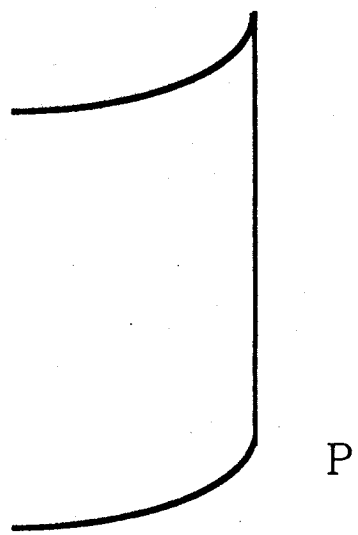


Fig. 4.3 P is a tangent continuity but a curvature discontinuity.

apex, which is a degenerate line of curvature, is included. If there are more than 4 segments along the boundary, then the problem is whether or not the surface can be divided into smaller four-segment surface patches by adding lines of curvature.

4.3 A net-knitting algorithm

Lines of curvature form two mutually orthogonal families of curves that cover the surface simply and without gaps like a piece of cloth. If sampled sparsely, it is like a net, both two-dimensionally and three-dimensionally. Having segmented a closed boundary into four segments, we knit a line-of-curvature net inside the given boundary. Since how to knit the net depends on how the lines of curvature flow over the surface, we first present a theorem that states the property of constant ratio intersection of lines of curvature over a class of surfaces, and then propose an algorithm for the net-knitting in image. Finally experimental examples and discussions are also given.

A. Theorem of Constant Ratio Intersection

Before we present the theorem, we first fix notations and give preparatory explanations.

As shown in Fig. 4.4, a surface patch bounded by four lines of curvature can be expressed by using the lines of curvature as the parametric lines in the form

$$\mathbf{x} = \mathbf{x}(u, v), \quad u1 \leq u \leq u2, \;\; v1 \leq v \leq v2, \tag{4.1}$$

where $\mathbf{x}$ is $(x, y, z)^T$. The four intersections are P, $\mathbf{x}(u1, v1)$, Q, $\mathbf{x}(u2, v1)$, R, $\mathbf{x}(u2, v2)$ and S, $\mathbf{x}(u1, v2)$, and the four lines of curvature, PQ, QR, SR and PS are

$$\begin{aligned}
\mathbf{x} &= \mathbf{x}(u, v1), \quad u1 \leq u \leq u2, \\
\mathbf{x} &= \mathbf{x}(u2, v), \quad v1 \leq v \leq v2, \\
\mathbf{x} &= \mathbf{x}(u, v2), \quad u1 \leq u \leq u2, \\
\mathbf{x} &= \mathbf{x}(u1, v), \quad v1 \leq v \leq v2.
\end{aligned}$$

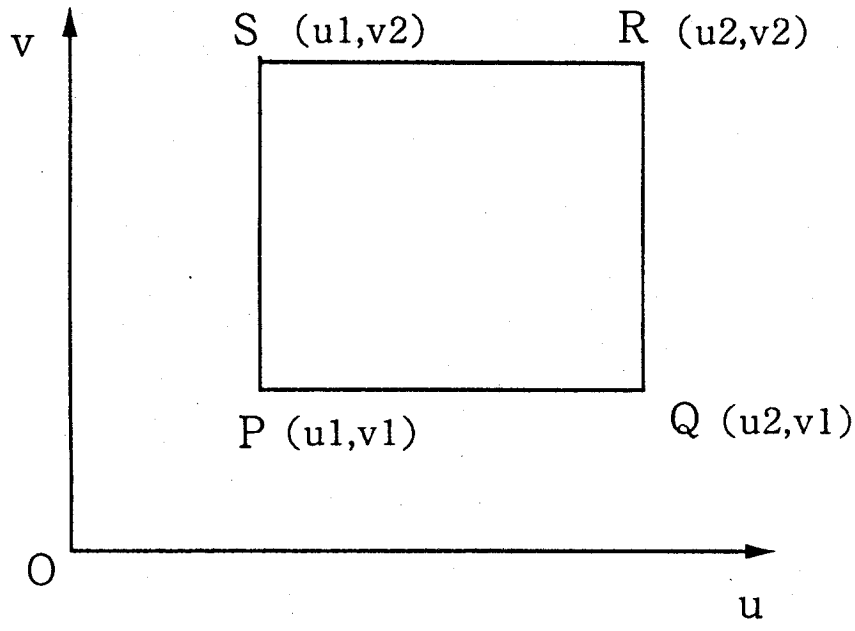


Fig. 4.4 Parameterized coordinate system

By *constant ratio intersection* of a line of curvature u = const., we mean

$$\frac{d}{dv}\left(\frac{\int_{u1}^{u1+k(u2-u1)} \sqrt{\frac{dx}{du}\cdot\frac{dx}{du}}\,du}{\int_{u1}^{u2} \sqrt{\frac{dx}{du}\cdot\frac{dx}{du}}\,du}\right) = 0, \quad 0 < k < 1; \tag{4.2}$$

i.e., the line of curvature $u = u1 + k(u2 - u1)$ divides the arc length of any line of curvature v = const. that it intersects at the same ratio.

[Theorem] (1) If the surface patch (that is bounded by four lines of curvature and without umbilics on it) is taken from a surface of revolution, then a meridian intersects any parallel at a constant ratio, and a parallel intersects any meridian at a constant ratio inside the boundary; (2) If the surface patch is taken from a generalized cone whose axis is straight and perpendicular to the planar cross section, then a fluting intersects any skeleton at a constant ratio inside the boundary; and (3) If the surface patch is taken from a developable surface whose Gaussian curvature is zero, then a line of curvature other than a ruling intersects any ruling at a constant ratio inside the boundary. [end]

The proof of this theorem is given in the appendix.

B. Algorithm

We first describe the algorithm and then give explanations.

As shown in Fig. 4.5, the segments a0-a1, b0-b1, c0-c1 and d0-d1 bound a surface patch. The points a2, b2, c2 and d2 are the center points that divide each segment into equal chord lengths. The points a3, a4, b3, b4, c3, c4, d3 and d4 are the chord centers of the new segments. We can go further until sufficient resolution is reached. We first find the point ab2cd2, which has both an equal distance to the points a2 and b2, and an equal distance to the points c2 and d2. The point is regarded as the intersection of the segments a2-b2 and c2-d2. Similarly, we can find the intersection of the segments a2-ab2cd2 and c3-d3. Repeating this process, we have two point sets that are dense enough to approximate the segments a2-b2 and c2-d2. Interestingly, they divide the

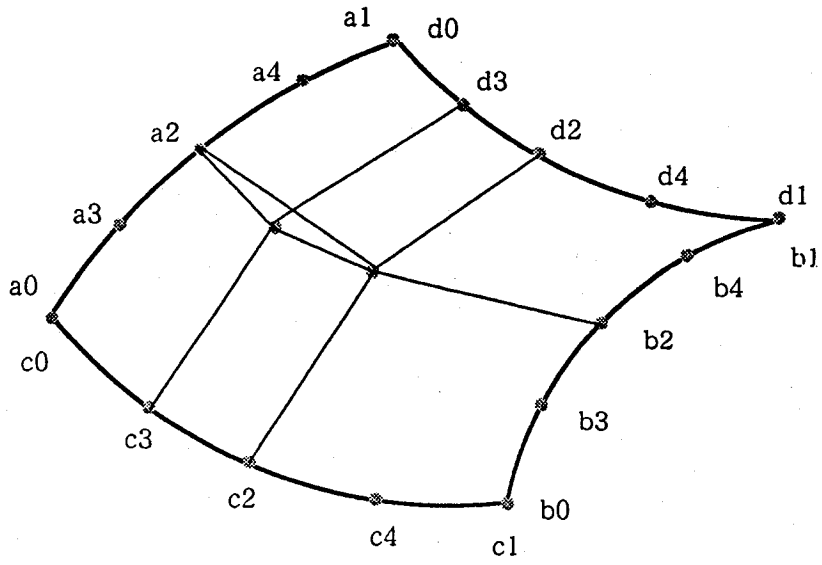


Fig. 4.5 A net-knitting algorithm

the surface into four subsurfaces, to which the algorithm described above can be recursively applied to knit a dense LOC net.

The algorithm is based on three assumptions: (1) both the lines of greatest curvature and the lines of least curvature have the property of constant ratio intersection; (2) the arc center of each segment in image is the projection of the arc center of the corresponding segment in space; and (3) the arc center of each segment has the same distance to the two end points.

If the assumption (1) holds, then we can draw a line of curvature through the arc centers of the boundary segments (in space) of each pair. As stated in the theorem, a broad class of surfaces possesses the property of constant ratio intersection of lines of curvature. Empirically, if the boundary segments of each pair are similar, the lines of curvature on surfaces which do not strictly satisfy the assumption (1) can still be treated as they do.

If the assumption (2) holds, then the segments drawn through the arc centers of the boundary segments of each pair are the projections of the lines of curvature through the arc centers of the corresponding segments in space. Because of the foreshortening, this assumption does not always hold. However, if the space segments of each pair are foreshortened in nearly the same way, then we can still draw lines of curvature through the arc centers of the image segments of each pair, though they are no longer the projections of the lines of curvature through the arc centers of the corresponding space segments.

If the assumption (3) holds, then we can find the arc centers of the lines of curvature inside the boundary before the complete segments are reconstructed. Provided that the segments are curved in a similar way, being based on this assumption does not introduce error, even if the assumption does not hold strictly.

In summary, if these assumptions do no hold, the obtained net only approximate the original surface. We call this net the first-order approximation. If,

however, the curvature of the boundary segments is not too great, then the error is within tolerance.

C. Experimental examples and discussions

The algorithm has been implememnted on a lisp machine. Two examples are shown in Fig. 4.6 and Fig. 4.7. The output nets are intuitively satisfactory. Note that the input boundary segments can be either extremal or discontinuity edges.

The algorithm works perfectly on cylindrical surfaces. Given two parallel straight line segments and two parallel curved segments as the boundary, the algorithm outputs an accurate line-of-curvatur net inside the boundary, which is identical with the net constructed by Stevens (1981). For cylindrical surfaces, the assumption (1) holds strictly, while the other two do not. The obtained nets are accurate because the boundary segments of each pair are completely the same. For an arbitrary surface, as discussed before, the net knit by this algorithm is not as accurate as for a cylindrical surface.

The first-order approximation can be refined. The assumption (3) can be removed by modifying the chord centers to be arc centers. This is done by a 2-D relaxation method. We call the net so modified the second-order approximation. The assumption (2) can also be removed. Since the 3-D arc length can be calculated by integration after surface orientation at the intersections is obtained, the 2-D arc centers can be modified to be 3-D arc centers again by relaxation. We call this modified net the third-order approximation.

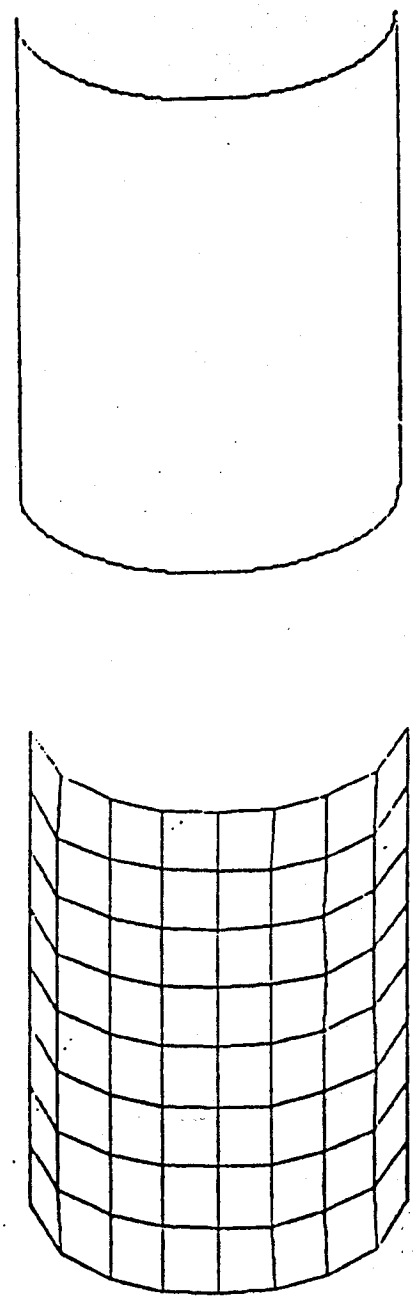


Fig. 4.6 A boundary (a) and its constructed net (b)

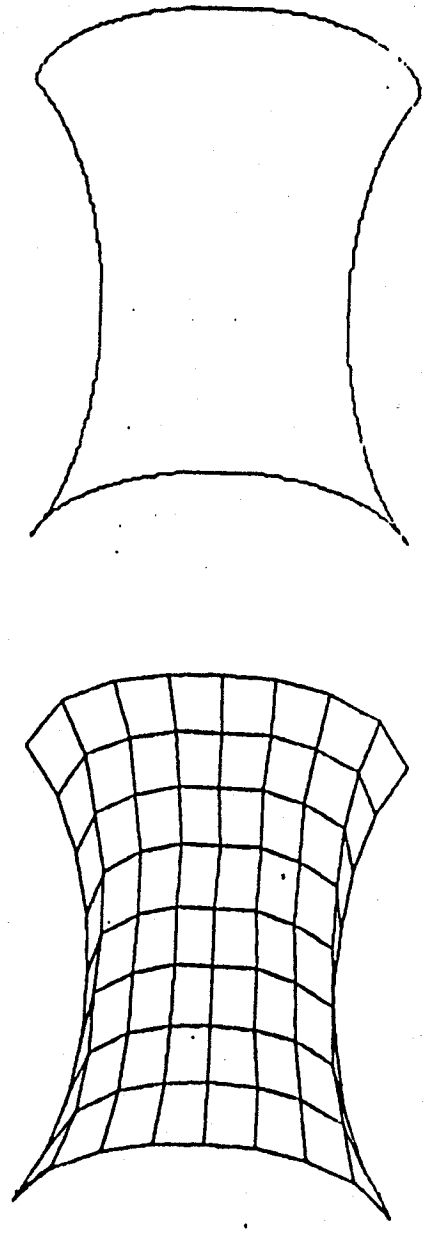


Fig. 4.7 A boundary (a) and its constructed net (b)

4.4 Surface orientation computation

A line-of-curvature net is a bridge between a 2-D boundary and its 3-D surface shape. Once the net is knit, the surface orientation at the intersections can be estimated. The lines of greatest curvature and the lines of least curvature intersect each other at a right angle in space, but they are foreshortened to obtuse angles (plus acute angles) in image. Stevens (1981) proposes the bisector method to estimate the surface orientation at each intersection, which is expressed by tilt and slant (see also [Witkin, 1981].) The more obtuse the angle, the more accurate the estimation. If the obtuse angle reaches two right angles, then the surface orientation is uniquely determined; the slant is a right angle, and the tilt is the normal of the image segment.

Stevens (1981) also proposes a method of propagating the surface orientation along the parallels on a cylindrical surface, from places where it is determined accurately to places where it is not. In image, a parallel intersects the straight rulings, which have a constant orientation, at different angles. As shown in Fig. 4.8, when the angle changes from $\beta 1$ to $\beta 2$, we have the equation

$$\tan \tau 1 \tan \beta 1 = \tan \tau 2 \tan \beta 2, \tag{4.3}$$

where $\tau 1$ and $\tau 2$ are the tilts. If $\tau 1$, $\beta 1$ and $\beta 2$ are known, then $\tau 2$ can be calculated.

However, for an arbitrary smooth surface, both the lines of greatest curvature and the lines of least curvature change their orientations. As shown in Fig. 4.9, suppose that when we move along a line of curvature from an intersection to the next one, V turns to V' and U turns to U'. Provided that the changes are not great, we can modify the method proposed by Stevens (1981) into a two-step approximation method. Without loss of generality, assume $\Delta\beta 2 > \Delta\beta 1$. We first fix U (let $\Delta\beta 1 = 0$) and calculate a new tilt due to only $\Delta\beta 2$. Next we fix V' and

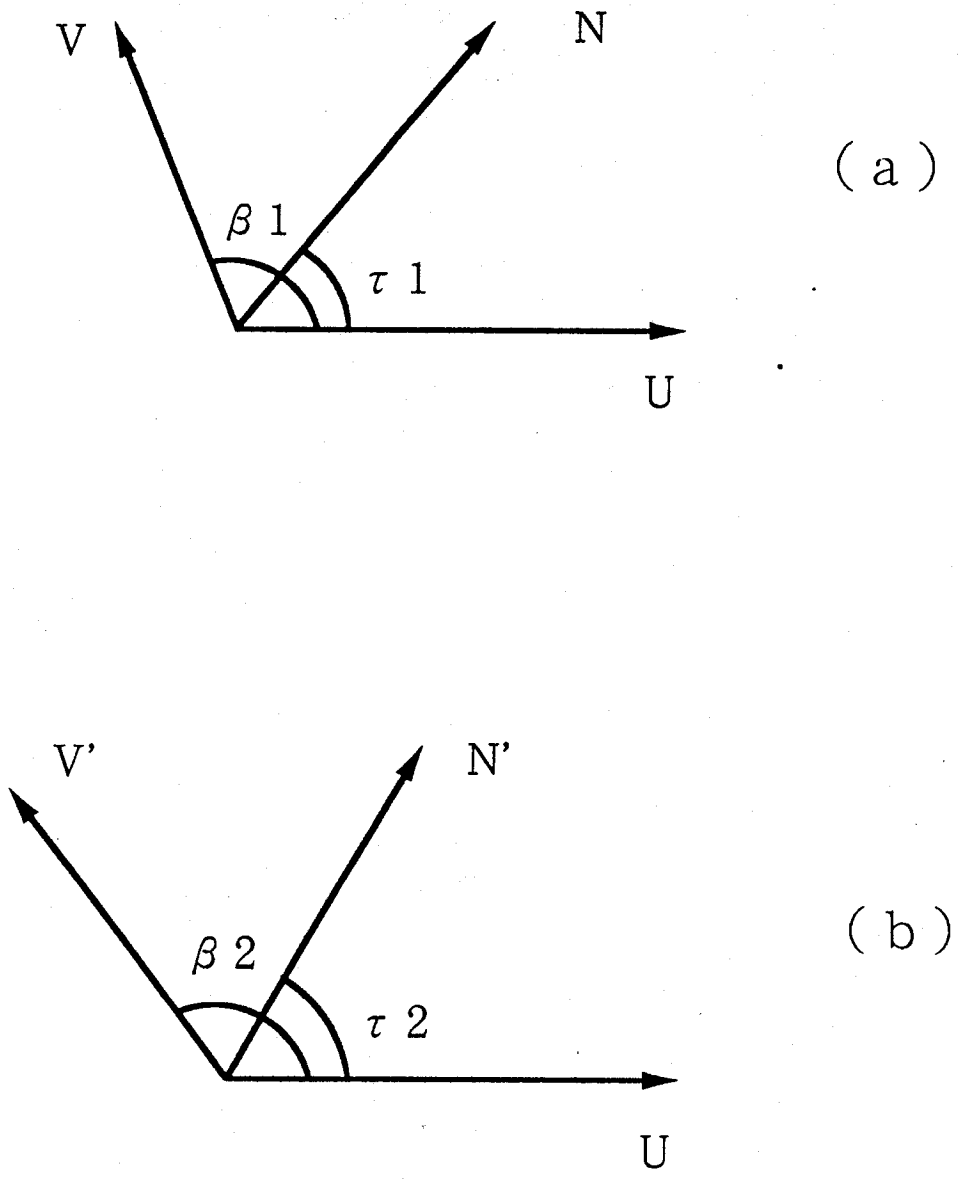


Fig. 4.8 Tilt variation from intersection angle variation on a cylindrical surface

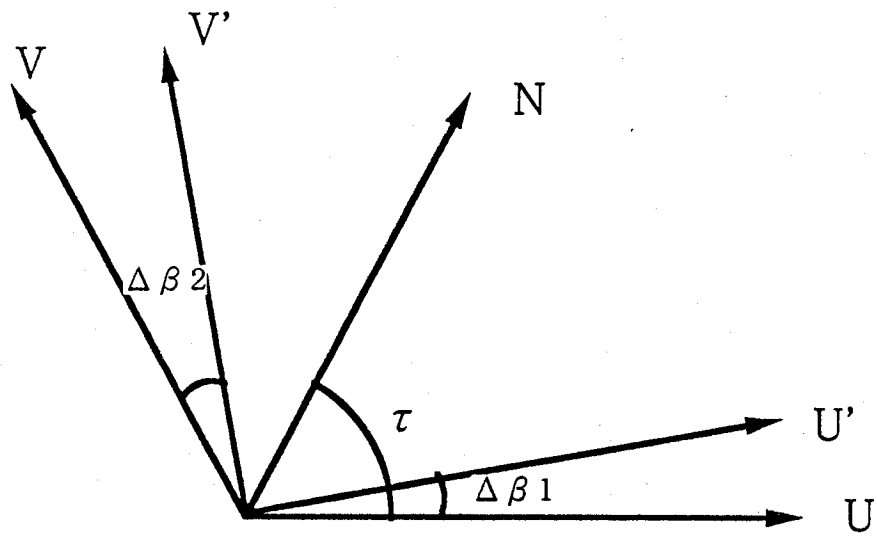


Fig. 4.9 Tilt variation from intersection angle variation on a doubly-curved surface

calculate a new tilt due to $\Delta\beta 1$. The obtained tilt is the required one. Once the tilt τ is determined, the slant σ can also be calculated as

$$\sigma = \tan^{-1} \sqrt{\frac{\tan \tau}{\tan^3 \beta}}. \tag{4.4}$$

In summary, after the surface orientation at the most obtuse angle is determined, we can propagate it along the lines of curvature to all the other intersections by the method just described.

Chapter 5 Interpreting Quadrilaterals under Gravity

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. In this chapter we first analyze the relation among the camera, the ground and the rectangles supported by the ground, and then derive constraints to determine the rectangle orientation.

5.1 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 (an exception is the space stations where objects are in a gravity-free state.) Unlike the rectangularity, this is a natural regularity objectively existing in the external world. Thus 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 many cases, 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 that 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 [Sedgwick & Levy, 1985] (see section 7.2 for a detailed discussion on the representation of visual information.)

5.2 Camera-ground model

Let the normal of the ground plane be expressed by n_g in the viewer-centered coordinate system. n_g actually implies the relation between the perceiver and the ground. As discussed in Section 2.3, to keep stability, the horizontal axis of the image plane is favorably 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 so that the optical axis of the camera points obliquely to the ground, as humans look some feet ahead on to the road. For a space vertical line at an arbitrary location to be projected as an image vertical line, it is necessary that the projection be orthographic. The camera-ground model is shown in Fig. 5.1. Assuming this camera-ground model, we have

$$n_g = (0, 1, -\tan \alpha). \tag{5.1}$$

Before concluding this section, we have two points to note. The first is that the orthographic projection is a necessity of the vertical-to-vertical mapping, but it at the same time does not exclude perspective projection for other purposes. The rectangles can still be projected as non-parallelograms. To perceive local shape of an object, the perspective information is required; whereas to perceive the more global relation between the perceiver and the ground, the orthography is required. It is really of interest to observe human's this flexibilty to swing between orthographic and perspective projections.

The second is that the camera model is employed to perceive objects lying on the ground and thus is not suitable to perceive objects suspending from a ceiling. However, if we assume a minus α, then many of the properties can be inherited.

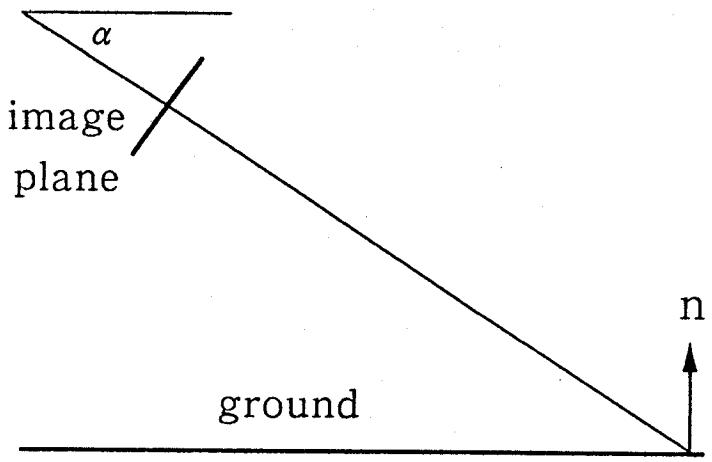


Fig. 5.1 The camera-ground model

5.3 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 behind it. When only one of the corners contacts the ground, it is most likely that we perceive the rectangle standing vertically, if some condition is met.

Which contact relation is perceived is largely dependent 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 (Fig. 5.2a) — i.e., the rectangle is remarkably slanted towards the sky — because we are not used to looking straight downward (see Chapter 4 and [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 only if one of the diagonal, of which the midpoint is the centroid, is vertical in image, and the upper and lower corners are acute angles (Fig. 5.2b) — again because of the posture assumed by the perceiver. The one side contact relation is perceived if the full contact and the one corner contact 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 (Fig. 5.2c,d).

If the perspective projection is assumed, a rectangle is no longer projected

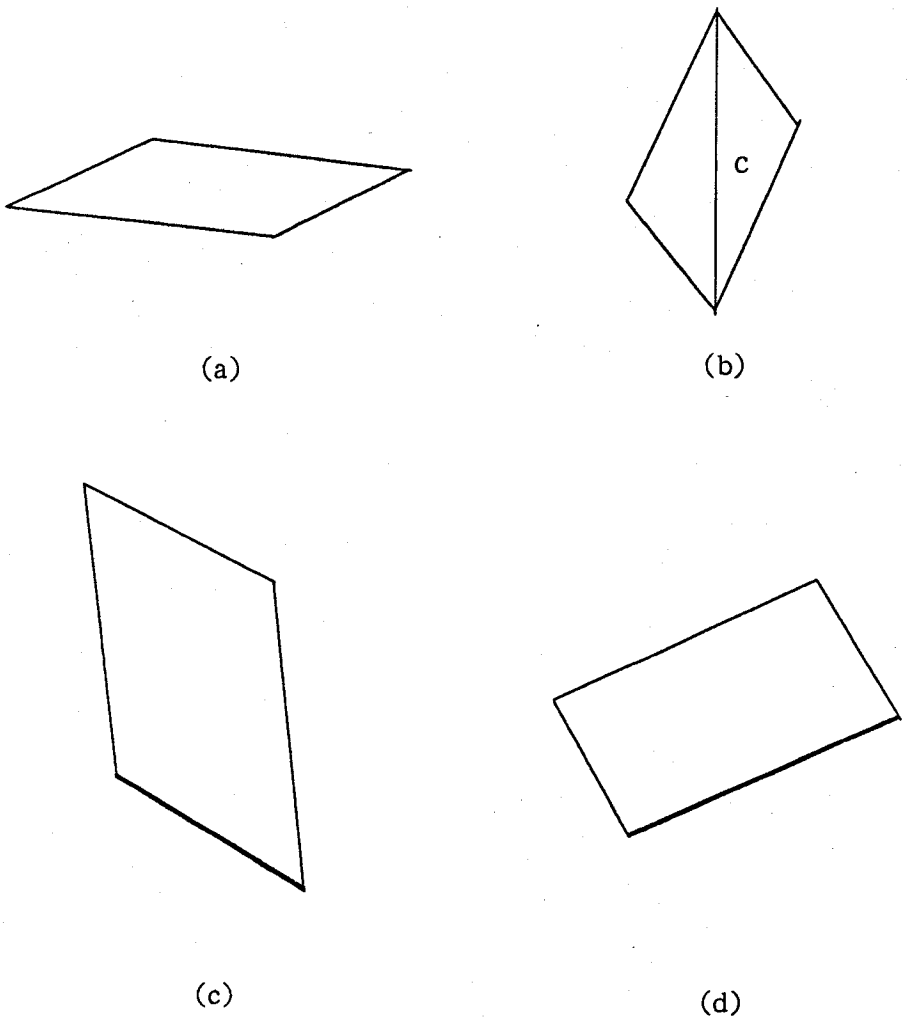


Fig. 5.2 The contact relations

as a parallelogram, but as a generic quadrilateral. In this case, there are two vanishing points. If the rectangle lies on the ground, then the two vanishing points must lie on the horizon (Fig. 5.3).

We do not intend to claim the completeness of the analysis, because perception of one contact relation is not necessarily exclusive of another and the conditions are not completely quantified. Even so, however, if we manage somehow to quantify the conditions — e.g., in the case of 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 (see Section 7.2.3 for a discussion on this subject.)

The above three contact relations can be respectively expressed as

$$\mathbf{PF} \cdot \mathbf{n_g} = 0, \quad \text{and} \quad \mathbf{QF} \cdot \mathbf{n_g} = 0; \tag{5.2}$$

$$(\mathbf{PF} \times \mathbf{QF}) \cdot \mathbf{n_g} = 0; \tag{5.3}$$

$$\mathbf{PF} \cdot \mathbf{n_g} = 0, \quad \text{or} \quad \mathbf{QF} \cdot \mathbf{n_g} = 0. \tag{5.4}$$

In the following we discuss two special cases in which one pair of edges of the parallelogram is either vertical or horizontal.

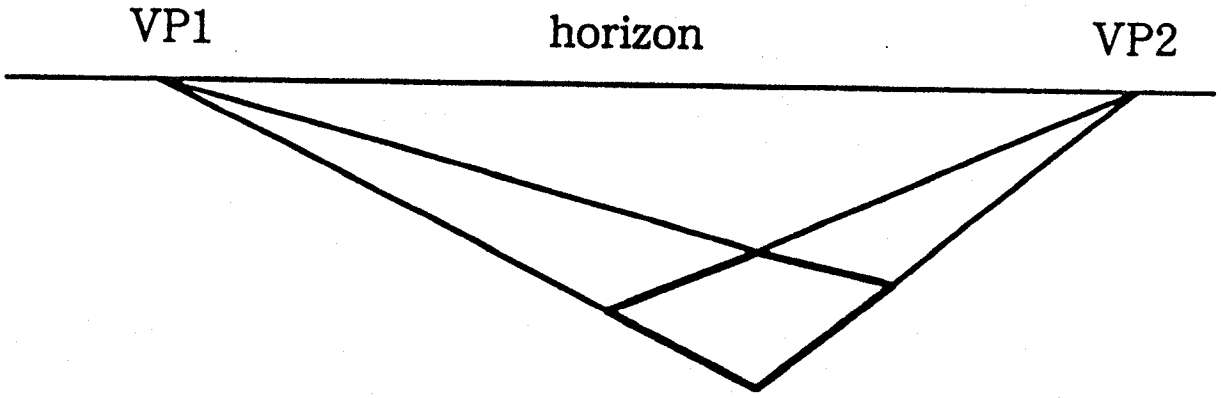


Fig. 5.3 If the rectangle lies on the ground, then the two vanishing points must lie on the horizon.

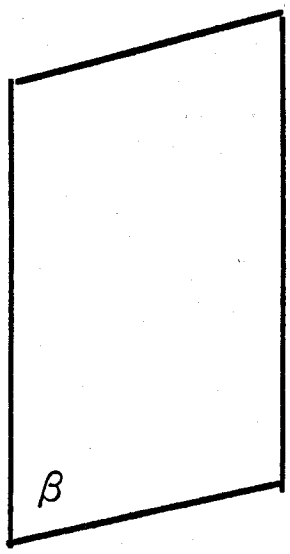


Fig. 5.4 A rectangle standing vertically on the ground

(Case1)

Since it is a parallelogram, the projection is orthographic. Without loss of generality, suppose that $\mathbf{PO} = (0, 1, c)$ is vertical, and $\mathbf{QO} = (1, b', c')$ is not vertical (Fig. 5.4). If the image angle between PO and QO is β ($\beta \neq 0$), then

$$\mathbf{PO} \cdot \mathbf{QO} = (0, 1) \cdot (1, b') = |(0, 1)||(1, b')| \cos \beta;$$

i.e.,

$$b' = \cot \beta, \tag{5.5}$$

PF is interpreted as a vertical in space and thus is identical with $\mathbf{n_g}$. Since PF is vertical, QF must be parallel to the ground. Thus we have

$$\mathbf{QF} \cdot \mathbf{n_g} = (1, \cot \beta, c') \cdot (0, 1, -\tan \alpha) = 0;$$

i.e.,

$$c' = \cot \alpha \cot \beta. \tag{5.6}$$

The normal of the rectangle is then

$$\mathbf{n} = \mathbf{PF} \times \mathbf{QF} = \mathbf{n_g} \times \mathbf{QF} = (0, 1, -\tan \alpha) \times (1, \cot \beta, \cot \alpha \cot \beta). \tag{5.7}$$

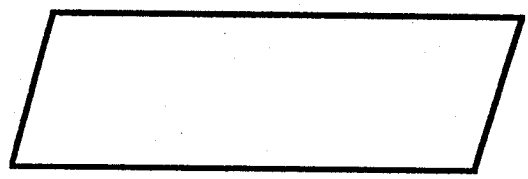


Fig. 5.5 If the horizontal lines are interpreted as parallel to the image plane, the parallelogram cannot be interpreted as a rectangle. But it is perceived by our eyes as a rectangle obliquely standing on the ground.

(Case2)

Without loss of generality, suppose that the image of PF is horizontal, and that of QF is not. As discussed in Section 2.3, PF is horizontal and parallel to the image plane, expressed by $\mathbf{PF} = (1, 0, 0)$. Since

$$\mathbf{PF} \cdot \mathbf{QF} = 0,$$

and

$$\mathbf{PF} \cdot \mathbf{n_g} = 0,$$

we have

$$\mathbf{PF} \cdot (\mathbf{QF} - \mathbf{n_g}) = 0. \tag{5.8}$$

That is, QF must also be vertical. Thus the parallelogram shown in Fig. 5.5 cannot be interpreted as a rectangle. This contradicts the perception by human eyes. In fact we perceive the figure as a rectangle standing obliquely, not vertically, on the ground. This is another example of the difference between interpretations at the geometrical and perceptual levels (see Chapter 6.)

Chapter 6 Quadrilaterals as Faces of a Rectangular Polyhedron

Any figures can be interpreted at two different levels: the geometrical level and the perceptual level [Sugihara, 1986]. While interpretations at the geometrical level must strictly obey mathematics, interpretations at the perceptual level show humanlike flexibility. Interpretations at two levels may be fairly different. We first deal with the geometrical level.

6.1 The geometrical level

It can be mathematically proven that a trihedral polyhedron with quadrilateral faces is a hexahedron. If all the faces are parallelograms, the hexahedron becomes a parallelepiped. If the parallelograms are rectangles, then the parallelepiped becomes a rectangular polyhedron.

Trivially, when a parallelepiped is projected onto an image from a general position, orthographically or perspectively, at most three of its six faces are visible. All the three visible faces in the image are parallelograms if the projection is orthographic, and are not if the projection is perspective.

Now we consider the inverse problem — how to infer the original object from its non-degenerate image [Kender & Freudenstein, 1987]. We are given three quadrilaterals, every two of which have a common edge — this condition is sufficient to define their interrelations. Since there are three common edges, the total number of edges is 9 (=12-3). They can be grouped into three, each of which has three non-intersecting edges, as *e1-e2-e3, e4-e5-e6, e7-e8-e9* shown in Fig. 6.1. The first condition for the figure to mean a real object is that the edges of each group, when extended, meet at a common point, which may approach infinity.

In the following we try to find the conditions for the three quadrilaterals to be simultaneously interpreted as rectangles.

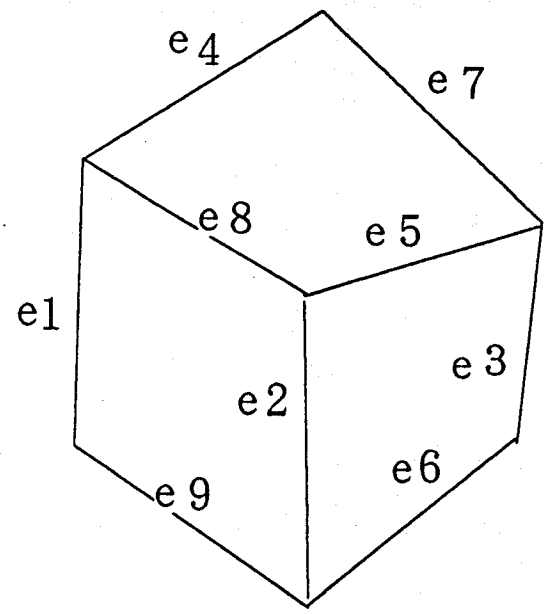


Fig. 6.1 Three quadrilaterals have totally 9 edges, if every two quadrilaterals have a common edge.

If none of the three intersection points approaches infinity, we need only to examine whether or not the three hemispheres, each of which is determined by the two vanishing points (the common intersection points) as described in Section 3.2, have a common point; i.e., whether or not the following three equations have a solution for $(x0, y0, f)$.

$$\begin{aligned}
&(x1 - x0)(x2 - x0) + (y1 - y0)(y2 - y0) + f^2 = 0,\\
&(x2 - x0)(x3 - x0) + (y2 - y0)(y3 - y0) + f^2 = 0,\\
&(x3 - x0)(x1 - x0) + (y3 - y0)(y1 - y0) + f^2 = 0,
\end{aligned} \tag{6.1}$$

where $(x1, y1)$, $(x2, y2)$ and $(x3, y3)$ are the three intersection points. Taking the differences of every two equations, we have

$$\begin{aligned}
&(x1 - x0)(x3 - x2) + (y1 - y0)(y3 - y2) = 0,\\
&(x2 - x0)(x1 - x3) + (y2 - y0)(y1 - y3) = 0,\\
&(x3 - x0)(x2 - x1) + (y3 - y0)(y2 - y1) = 0.
\end{aligned} \tag{6.2}$$

These equations mean that $(x0, y0)$ is the orthocenter of the triangle (Fig. 6.2) formed by the three intersection points $(x1, y1)$, $(x2, y2)$ and $(x3, y3)$. Thus, fortunately, the three hemispheres always have a common point.

If only one of the three intersection points approaches infinity, then we have the following three equations,

$$\begin{aligned}
&a(x1 - x0) + b(y1 - y0) = 0,\\
&a(x2 - x0) + b(y2 - y0) = 0,\\
&(x1 - x0)(x2 - x0) + (y1 - y0)(y2 - y0) + f^2 = 0,
\end{aligned} \tag{6.3}$$

where (a, b) is the orientation vector, and $(x1, y1)$ and $(x2, y2)$ are the coordinates of the other two intersection points. The difference of the first two equations means that the parallel edges are perpendicular to the line linking the other two

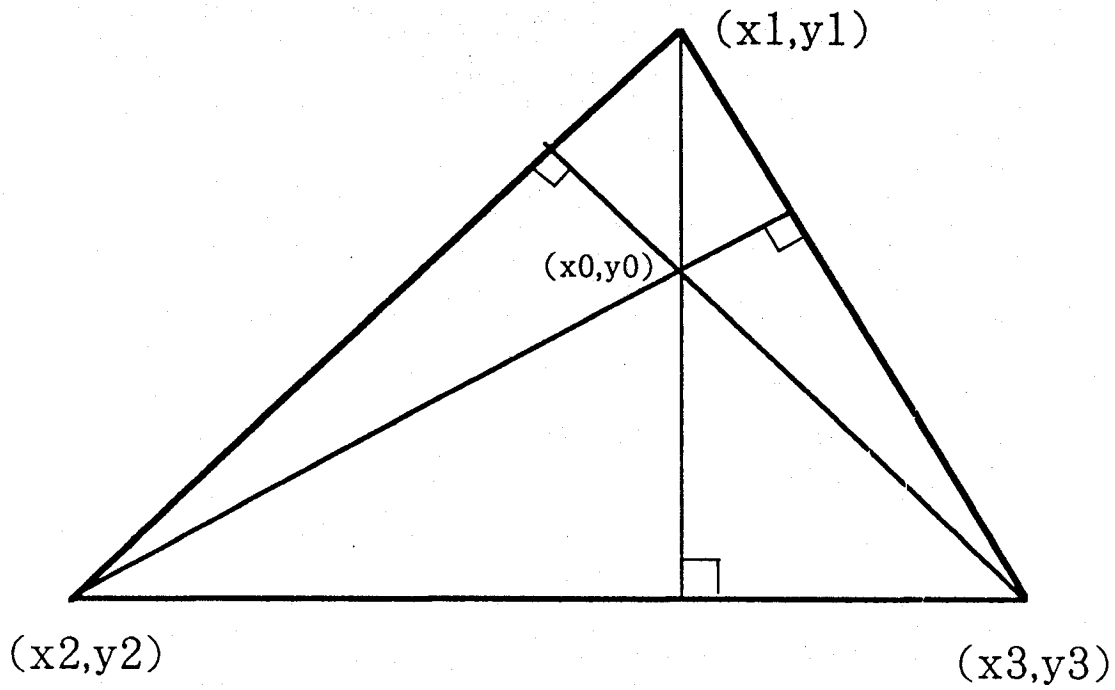


Fig. 6.2 The orthocenter of the triangle is the "image center".

intersection points. The focal point is not completely determined, but constrained to lie on a semicircle, whose two end points are the two intersection points, and the plane on which the semicircle lies is perpendicular to the image plane (Fig. 6.3).

If two of the three intersection points approach infinity and the other one does not, then one of the three quadrilaterals must be a rectangle. As discussed in Section 3.2, if two vanishing points approach infinity, then the corresponding image quadrilateral is a parallelogram or a rectangle. If the parallelogram is to be interpreted as a rectangle in space, then the projection must be orthographic. If the projection is orthographic, then all rectangles in space are projected as parallelograms in image. Thus, if two of the three intersection points approach infinity and their corresponding quadrilateral is not a rectangle, then the other one must also approach infinity. As described in Section 3.2, if a space rectangle is parallel to the image plane, then it is projected as an image rectangle under the perspective projection. If it is a face of a rectangular polyhedron, then the other two visible faces are projected as generic quadrilaterals, with the corresponding vanishing points not approaching infinity.

If all the three intersection points approach infinity, then the projection is orthographic. The condition for a rectangular polyhedron interpretation is that the following three equtions have a common solution for $c1$, $c2$ and $c3$.

$$\begin{aligned} \mathrm{V1}\cdot\mathrm{V2} + c1c2 &= 0, \\ \mathrm{V2}\cdot\mathrm{V3} + c2c3 &= 0, \\ \mathrm{V3}\cdot\mathrm{V1} + c3c1 &= 0, \end{aligned} \tag{6.4}$$

where $\mathrm{V1} = (a1, b1)$, $\mathrm{V2} = (a2, b2)$ and $\mathrm{V3} = (a3, b3)$ are the orientation vectors of the three groups of parallel edges. Easily we have

$$c1^2 = -\frac{(\mathrm{V1}\cdot\mathrm{V2})(\mathrm{V1}\cdot\mathrm{V3})}{(\mathrm{V2}\cdot\mathrm{V3})},$$
$$c2^2 = -\frac{(\mathrm{V2}\cdot\mathrm{V3})(\mathrm{V2}\cdot\mathrm{V1})}{(\mathrm{V3}\cdot\mathrm{V1})},$$

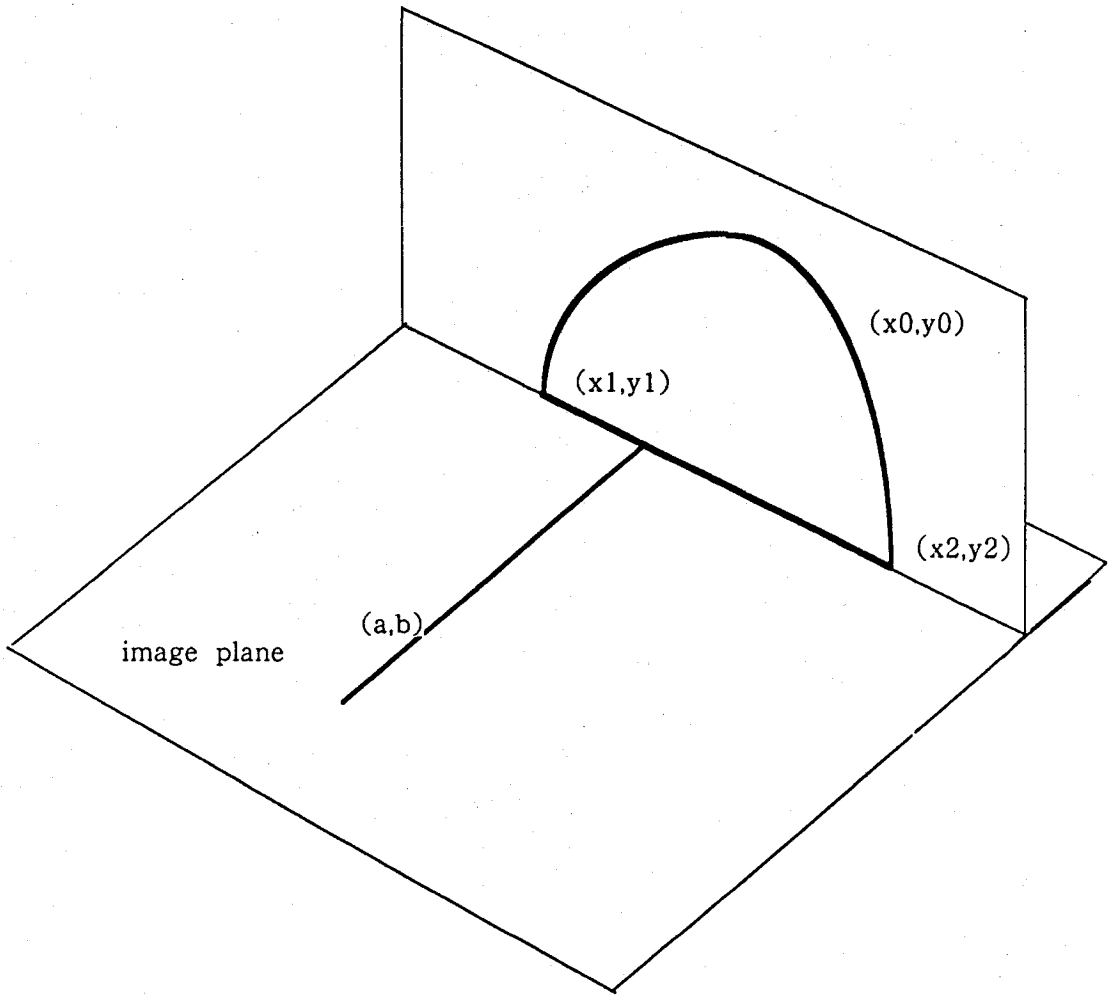


Fig. 6.3 The focal point is constrained to lie on the semicircle.

$$c3^2 = -\frac{(V3 \cdot V1)(V3 \cdot V2)}{(V1 \cdot V2)}. \qquad (6.5)$$

The condition for a solution of $c1$, $c2$ and $c3$ is that all the three inner products are negative — all the three angles around the central corner are greater than 90 degrees. It is impossible that two of the three inner products are positive and the other one negative, because otherwise one of the angles around the central corner would be greater than 180 degrees, reducing the three quadrilaterals to two.

6.2 The perceptual level

While interpreting figures at the geometrical level is strictly governed by mathematics, it must possess certain degree of flexibility at the perceptual level; otherwise it would not work on hand-drawn figures in a humanlike way [Sugihara, 1986]. Both the condition for a real object interpretation and the condition for a rectangular polyhedron interpretation derived at the geometrical level have te be changed. First, in a hand-drawn figure, it is too strict a condition that the three edges of each group, when extended, meet at a common point. Also, the parallel edges are only nearly parallel. Secondly, as given in Fig. 6.4, even when one of the three angles around the central corner is exactly 90 degrees, the figure is still perceived as a rectangular polyhedron (the interpretation is mathematically impossible.) A feasible explanation for this fact is that the individual quadrilaterals are first interpreted separately (as rectangles) and then integrated as faces of a rectangular polyhedron in a less strict way than at the geometrical level. This kind of perception seems quite ubiquitous. When drawing a man or an animal, children usually put together a frontal view of the face and a side view of the body. But the flexibility has its limit; if one of the three angles around the central corner is less than 90 degrees, the figure is no longer perceived as a rectangular polyhedron (Fig. 6.5,) but an ordinary parallelepiped [Perkins, 1983; Kanade & Kender, 1983].

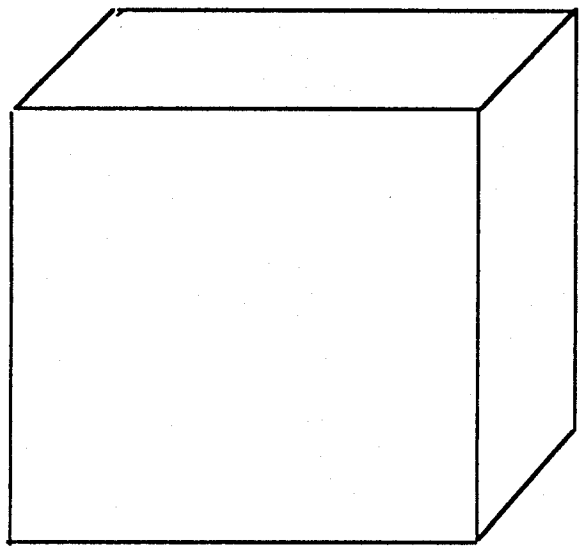


Fig. 6.4 Perception says, "it is a cube". Mathemetics argues, "you are wrong".

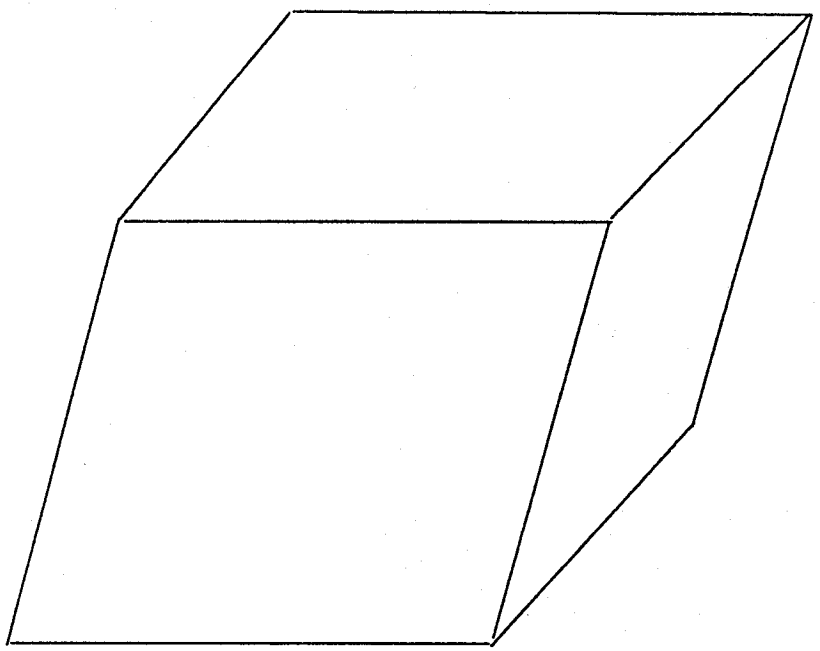


Fig. 6.5 If one of the three angles around the central corner is less than 90 degrees, the figure is no longer perceived as a rectangular poyhedron.

Lastly, although the rectangular polyhedron interpretation is mathematically correct if all the three intersection points of each edge group do not approach infinity, it does not always agree with human perception, if the orthocenter of the triangle formed by the three intersection points is far away from the figure itself, simply because we do not draw such figures to represent and communicate something.

Chapter 7 Summary and Discussions

We are now in a position to look back over the work described in the previous chapters, point out remaining problems and suggest directions for future research.

7.1 Summary

Throughout this thesis the main goal has been to develop computational theories for interpreting various image quadrilaterals. To develop specific algorithms to calculate solutions to the theories has only been the secondary goal. The strength of the computational approach lies in providing a framework within which mathematicians, computer scientists, psycophysicists and neurophysiologists can describe and communicate their understanding of the problems.

In the previous chapters we have introduced our philosophy of vision research, defined the goal of our research, contrasted our work with the background, and proposed several regularities and constraints to be employed in the visual interpretation of quadrilaterals. By studying these seemingly simple figures, we have obtained a number of insights into the nature of the general problem of interpreting image contours.

Following a general introduction in Chapter 1, Chapter 2 introduced basic terms and concepts, restated the *general viewpoint assumption* in our new framework, and then applied it to inferring 3-dimensional continuity and discontinuity, and to interpreting image vertical and horizontal lines. In Chapter 3, we proposed the rectangularity regularity to be the prime constraint in the visual interpretation of quadrilaterals. The nature of the regularity has been analyzed, and we believe that this kind of regularities is specific forms of our perceptual system's preference of regular forms over irregular forms, and should be called *subjective regularities*. By incorporating the rectangularity regularity, quadrilaterals in image are interpreted as rectangles in space, and the "image center" and focal length

are determined together with rectangle orientation. The orthographic projection is treated as the limit of perspective projection as f approaches infinity. In Chapter 4, we adapted the rectangularity regularity to apply to interpreting "curved quadrilaterals". Applying the line of curvature (LOC) regularity, a 4-segment closed image contour is interpreted as four 3-D lines of curvature on the surface that the image contour depicts. We further proposed a segmentation rule, an algorithm for constructing a net of lines of curvature inside the boundary, and a method for estimating surface orientations at the net intersections, and finally presented some experimental results. In Chapter 5, we discussed the role that gravity plays in visual perception and in interpreting quadrilaterals in particular. 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. We first analyzed the relation among the perceiver (camera), the ground and the rectangles supported by the ground, and then derived constraints to determine the rectangle orientation. In Chapter 6, we examined the conditions for interpreting quadrilaterals as faces of a rectangular polyhedron. There are two levels of interpretation: the geometrical level and the perceptual level. While interpretations at the the geometrical level must strictly obey mathematics, interpretations at the perceptual level show humanlike flexibility. Interpretations at the two levels are usually fairly different.

7.2 Representation of visual information

It is natural to write any descriptions of spatial relations and 3-D models into the image, which we coin as *visual map*, though the coexistence of 2-dimensionality and 3-dimensionality makes it sound paradoxical. Spatial relations can be described in both, or either of, the viewer-centered coordinate system [Marr, 1982] (when, e.g., it is used for navigation,) and a more objective world-centered coordinate system [Sedgwick & Levy, 1985]. Shape of objects is usually, if not exclusively, object-centered, so that mental rotations can be performed if needed. The point is that we can never imagine an absolute object-centered 3-D model, because a subject is always present, and he must assume a position and an orientation relative to the object. No reference frame other than the image, or the *visual map*, then, is more natural and straightforward. In this sense the process of visual perception can be understood as one of adding 3-D descriptions to the 2-D image, literally one of image interpretation.

7.3 Conditions for the rectangle interpretation

In Chapter 3 we have proposed the rectangularity regularity to be the prime constraint in the visual interpretation of quadrilaterals, by which quadrilaterals in image are all interpreted as rectangles in space. It is true that humans do tend to interpret quadrilaterals in image as rectangles in space, but it is also true that humans do not interpret all quadrilaterals in image as rectangles in space. We did not provide the conditions for a quadrilateral in image to be firmly interpreted as a rectangle in space in this thesis. However, we consider that the quantification of the conditions is possible through psycological experiments. It is our conviction that the parameters underlying the perception are the maximal possible focal length and the distance between the intersection of the two diagonal (that is the centroid if it is a rectangle) and the associated "image center". The longer the focal length, the shorter the distance from the intersection to the associated

"image center", the easier the quadrilateral is interpreted as a rectangle.

7.4 Perception of the contact relations

In Chapter 5, we analyzed the three contact relations between the ground and a rectangle supported by it. Although we can intuitively determine the contact relation when given a specific figure, the boundary between different interpretations is not always clear. If we are to incorporate this into a practical artificial vision system, it is required to completely quantify the conditions for each contact relation. Again psycological experiments can be conducted. Two parameters are sufficient to describe the "space" if a parallelogram rather than a generic quadrilateral is used as the object. One of the parameters is the lowest angle of the parallelogram, whose supplementary angle uniquely determines the other pair of angles, and the other parameter is the right-lower edge's angle to the horizontal axis. Varying the two parameters and asking the subjects to answer which contact relation they perceive, we can divide the 2-D "space" into three regions, each of which corresponds to a specific contact relation.

7.5 Towards a rule-based system

The regularity-based approach to the visual interpretation of line drawings opens an attractive route to the development of a rule-based system of line drawing understanding. Essentially, line drawings are more symbolic than images and thus more suited to symbolic processing, to which the matured AI tool of knowledge-based systems applies. We have in this thesis investigated a number of pieces of knowledge that are employed in human perception of line drawings. But they account for only a tiny portion. Still much is yet to be searched for. As the knowledge increases, the problem of conflicts among the rules will emerge. One solution to the problem is to assign a priority value to each rule. The humanlike flexibility will not be achieved if there is not a powerful mechanism to coordinate the rules.

Appendix Proof of the Constant Ratio Intersection Theorem

[Proof]

(1) If the z-axis is the axis of revolution of the curve $z = f(v)$ (Fig. A.1), the resulting surface can be given as

$$\mathbf{x} = (v\cos u,\ v\sin u,\ f(v)). \tag{A.1}$$

The line of curvature are the meridians and the parallels, given by u =const. and v =const., respectively. For the meridians, the ratio of length

$$\frac{\int_{u1}^{u1+k(u2-u1)} \sqrt{\frac{d\mathbf{x}}{du}\cdot\frac{d\mathbf{x}}{du}}\,du}{\int_{u1}^{u2} \sqrt{\frac{d\mathbf{x}}{du}\cdot\frac{d\mathbf{x}}{du}}\,du} = k, \quad 0<k<1; \tag{A.2}$$

is independent of v. For the parallels, the interval

$$\int_{v1}^{v1+k(v2-v1)} \sqrt{\frac{d\mathbf{x}}{dv}\cdot\frac{d\mathbf{x}}{dv}}\,dv = \int_{v1}^{v1+k(v2-v1)} \sqrt{1+[f'(v)]^2}\,dv, \quad 0<k<1; \tag{A.3}$$

is itself independent of u.

(2) If the planar cross section curve is given by

$$\mathbf{x} = f(u),\ y = g(u), \tag{A.4}$$

the z-axis is taken as the v-axis of the generalized cylinder, and the expansion function is $h(v)$ (Fig. A.2), then the resulting surface can be expressed as

$$\mathbf{x} = (\,h(v)f(u),\ h(v)g(u),\ v). \tag{A.5}$$

The lines of curvature are the flutings and the skeletons, given by u =const. and v =const., respectively. The arc length

$$\int_{u1}^{u1+k(u2-u1)} \sqrt{\frac{d\mathbf{x}}{du}\cdot\frac{d\mathbf{x}}{du}}\,du = \int_{u1}^{u1+k(u2-u1)} \sqrt{[f'(u)]^2+[g'(u)]^2}\,du, \quad 0<k<1; \tag{A.6}$$

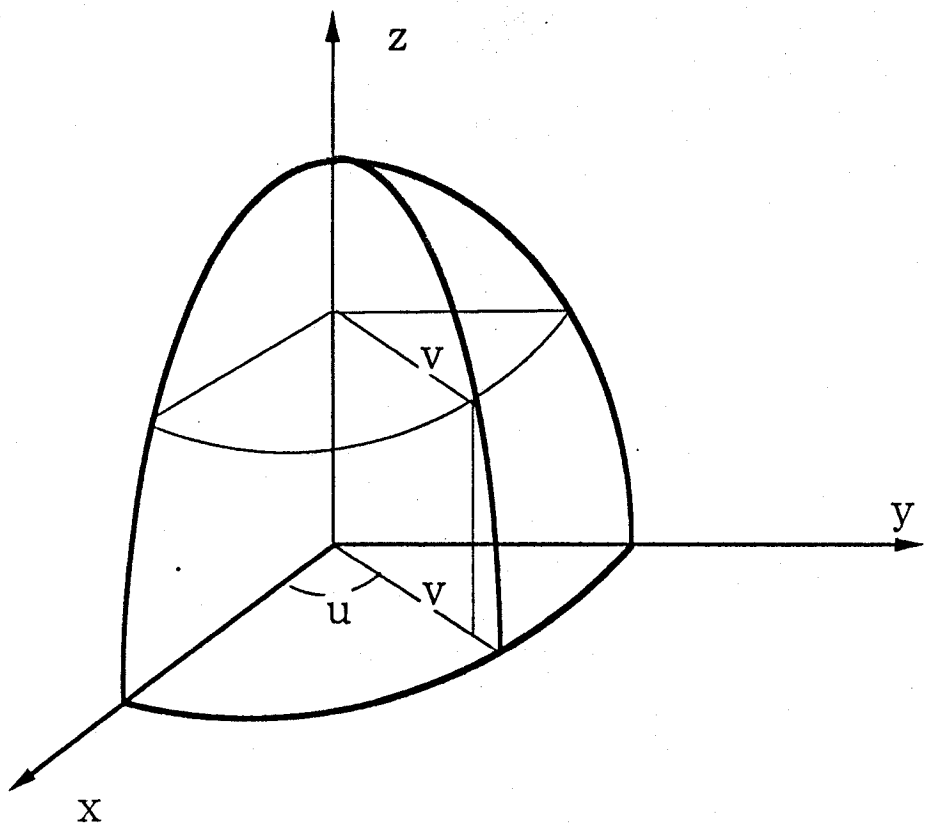


Fig. A.1 A surface of revolution

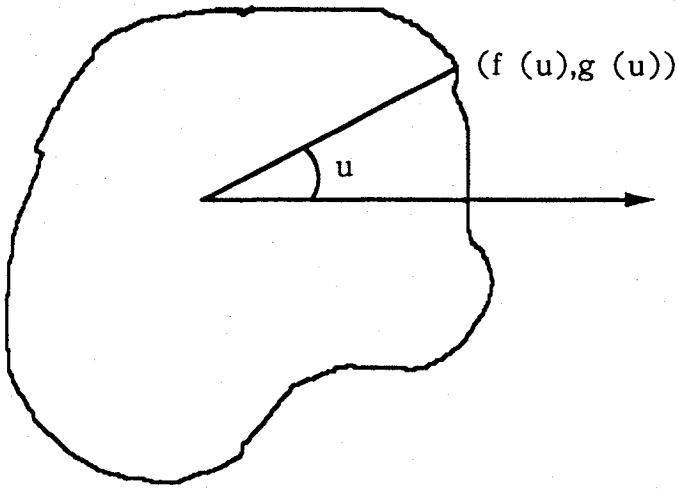


Fig. A.2 A cross section of a generalized cylinder

is itself independent of v.

(3) If a developable surface has zero Gaussian curvature, then one set of the lines of curvature are the straight lines. the surface can be rewritten as

$$\mathbf{x}(u,v) = \mathbf{A}(\mathbf{v})u + \mathbf{B}(\mathbf{v}). \qquad (A.7)$$

The straight lines are then given by u=const.. The length ratio

$$\frac{\int_{u1}^{u1+k(u2-u1)} \sqrt{\frac{d\mathbf{x}}{du}\cdot\frac{d\mathbf{x}}{du}}\,du}{\int_{u1}^{u2} \sqrt{\frac{d\mathbf{x}}{du}\cdot\frac{d\mathbf{x}}{du}}\,du} = \frac{\int_{u1}^{u1+k(u2-u1)} \sqrt{\mathbf{A}(v)\cdot\mathbf{A}(v)}\,du}{\int_{u1}^{u2} \sqrt{\mathbf{A}(v)\cdot\mathbf{A}(v)}\,du} = k, \qquad 0<k<1; \qquad (A.8)$$

is independent of v.

[end]

Bibliography

Barnard, S. (1983) Interpreting perspective images, Artificial Intelligence 21, pp. 435-462

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

Biederman, I. (1985) Human image understanding: Recent research and a theory, Computer Vision, Graphics and Image Processing, Vol. 32, pp. 29-73

Binford, T. (1981) Inferring surfaces from images, Artificial Intelligence 17 pp. 205-244

Brady, M. and Asada, H. (1984) Smoothed local symmetries and their implementation, International Journal of Robotics Research, Vol. 3, No. 3, pp. 36-61

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

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

Chakravarty, I. (1979) A generalizedline and junction labeling scheme with applications to scene analysis, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 1, pp. 202-205

Grimsom, W. (1981) From Images to Surfaces: A Computational Study of the Human Early Visual System, MIT Press

Hoffman, D. (1983) The interpretation of visual illusions, Scientific American 249, pp. 137-144

Hoffman, D. and Richards, W. (1986) Parts of recognition, From Pixels to Predicates, ed. by Pentland, A., Ablex

Huffman, D. (1971) Impossible objects as nonsense sentences, Machine Intelligence 6, ed. by Meltzer, B. and Michie, D., Edinburgh University Press

Kanade, T. (1980) A theory of origami world, Artificial Intelligence 13, pp. 279-311

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

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

Kanatani, K. (1986) The constraints on images of rectangular Polyhedra, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, pp. 456-463

Kanizsa, G. (1979) Organization in Vision: Essays on Gestalt Perception, Praeger

Kender, J. and Freudenstein, D. (1987) What is a "degenerate" view?, Proc. 10th Int. Joint Conf. on Artificial Intelligence, pp. 801-804

Koenderink, J. and van Doorn, A. (1982) The shape of smooth objects and the way contoors end, Perception, Vol. 11, pp. 129-137

Koenderink, J. (1984) What does the occluding contour tell us about solid shape? Perception, Vol. 13, pp. 321-330

Marr, D. (1982) Vision, W. H. Freeman

Marr, D. and Poggio, T. (1979) A theory of human stereo vision, Proceedings of the Royal Society (London) B204, pp. 301-328

Mackworth, A. (1973) Interpreting pictures of polyhedral scenes, Artificial Intelligence 4, pp.121-137

Palmer, S. (1983) The psycology of perceptual organization: A transformational approach, 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

Perkins, D. (1983) Why the human perceiver is a bad machine, Human and Machine Vision, ed. by Beck, Hope and Resonfeld, Academic Press

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 Computer Vision, pp. 662-666

Sedgwick, H. and Levy, S. (1985) Environment-centered and viewer-centered perception of surface orientation, Computer Vision, Graphics and Image Processing, Vol. 31, pp 248-260

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

Stevens, K. (1986) Inferring shape from contours across surfaces, From Pixels to Predicates, ed. by Pentland, A., Ablex

Struik, D. (1961) Differential Geometry, Addison-Wesley

Sugihara, K. (1979) Range-data analysis guided by a junction dictionary, Artificial Intelligence 12, pp. 41-69

Sugihara, K. (1986) Machine Interpretation of Line Drawings, MIT Press

Tsuji, S. Zheng, J. and Asada, M. (1987) Stereo vision of a mobile robot: World constraints for image matching and interpretation, Proc. IEEE Int. Conf. Robotics and Automation, pp. 1194-1199

Turner, K. (1974) Computer perception of curved objects using a television camera, Ph.D. Dissertation, Edinburgh University

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

Waltz, D. (1975) Understanding line drawings of scenes with shadows, The Psycology of Computer Vision, ed. by Winston, P., McGraw-Hill

Winston, P. (1984) Artificial Intelligence, Addison-Wesley

Witkin, A. (1981) Recovering surface shape and orientation from texture, Artificial Intelligence 17, pp. 17-46

Witkin, A. and Tenenbaum, J. (1983) On the role of structure in vision, Human and Machine Vision, ed. by Beck, Hope and Resonfeld, Academic Press

Witkin, A. and Tenenbaum, J. (1986) On perceptual organization, From Pixels to Predicates, ed. by Pentland, A., Ablex

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