Hypothesizing and Testing Geometric Attributes of Image Data

Kenichi Kanatani

Department of Computer Science, Gunma University
Kiryu, Gunma 376, Japan

Abstract

A general formalism for detecting geometric configurations of image data is presented. We first estimate an ideal geometric configuration that supposedly exists, and then check to what extent the original edges must be displaced *in order to support the hypothesis. All types of tests are reduced to computing a single* measure of edge displacement, *which provides a universal* measure of uncertainty *applicable to all types of decision-making.*

1. Introduction

In many computer vision problems, the 3D inference of images is made through two stages: the *image processing stage*, extracting geometric primitives such as edges and feature points, and the *computational stage*, computing such 3D properties as the shape and the position of the object. The accuracy of 3D information computed in the second stage depends largely on the accuracy of the first stage. However, there is one crucial difficulty that cannot be resolved by simply increasing the accuracy—the issue of *consistency* of the image data.

In many problems, a clue to 3D inference is provided by the information that image data, such as points and edges, have a special geometric configuration. For example, if several edges are *concurrent* (i.e., meeting at a common intersection when extended), we can infer that these edges may be parallel in the scene. However, inaccuracy due to noise and digitization is inevitable, and the required consistency is usually violated for real image data.

The problem we consider in this paper is the question of how we can detect geometric configurations of image data that observed image data apparently *do not* possess. In the past, many researchers have treated this problem with ad-hoc heuristics [1-3, 8], introducing measures to be thresholded and adjusting the threshold values. One possible mathematical basis is statistical analysis, but an appropriate statistical model is hard to obtain for real images, since image processing involves many operations in a complicated manner.

In this paper, we will present a theory based on the *hierarchical structure* of image processing without introducing any specific statistical model. Consider the following image processing scenario [3, 5]:

1. From a gray-level image, edges are detected.
2. Straight line segments are fitted to nearly straight edges.
3. Nearly collinear segments are replaced by a single line.
4. Vertex positions are computed as intersections of these lines.
5. 3D inferences are made from the line drawing thus obtained.

If we carefully look at these processes, we realize that there is a *hierarchy* of data. First, we obtain *primary data* by pixel-based image processing techniques, then *secondary data* by numerical computation over the primary data. It is reasonable to assume that *primary data consists of edges*, because edge detection is usually the first step of image processing; all subsequent data are derived from edges.

Our strategy is as follows. If we want to detect some geometric configuration on the image plane, we first *hypothesize* the configuration that supposedly exists, and then test *how much the original edges must be displaced* in order to support the hypothesis.

Although the threshold value must be empirically adjusted, the quantity to be thresholded is the amount of edge displacement alone. Since all image data are computed from the image data of lower levels, the computation can eventually be traced down to edges. This reduction to a *single criterion* serves as the *uncertainty level* of the decision, and we can compare the uncertainty levels of different types of decision with each other.

2. Mathematical Preliminaries

Given an image, define an xy-coordinate system on it such that the origin o *supposedly* indicates the position of the optical axis of the camera from which the image was obtained. Then, define an XYZ-coordinate system such that the Z-axis passes through the image origin o perpendicularly. The X- and Y-axes are taken to be parallel to the image x- and y-axes, respectively, and the origin O is taken to be in distance f from the image origin o, where f is *supposedly* the distance between the center of the lens and the surface of the film (often called the *focal length*) measured in the scale of the image coordinates.

This camera model *need not correspond to the true camera from which the image was obtained.* In other words, the camera model can be *hypothetical.* Although the choice of the camera model does not essentially affect the following arguments, we assume that the focal length f is approximately equal to the true value (at least having the same order).

The unit vector starting from the viewpoint O and pointing toward a point P in the image is called the *N-vector* of the point P, and the unit vector normal to the plane defined by the viewpoint O and a line l in the image is called the *N-vector* of the line l (see Fig. 1). From this definition. the N-vector

Part of this work was supported by the Ministry of Education, Science, and Culture under Grant in Aid for Scientific Research (No. 02805043).

of point (a,b) is $m=N[(a,b,f)^T]$, and the N-vector of line $Ax+By+C=0$ is be $m=\pm N[(A,B,C/f)^T]$ (the sign is arbitrary), where $N[u]=u/\|u\|$ denotes the normalization of vector u. Throughout this paper, we put $k=(0,0,1)^T$. It is easy to prove the following statements (We omit the proof).

Lemma 1. *If m is the N-vector of a point P in the image, the vector $\vec{OP}$ starting from the viewpoint O and ending at P is given by*

$$\vec{OP}=fm/(m,k). \tag{2.1}$$

Lemma 2. *If n is the N-vector of a line l in the image, the unit vector u perpendicular to the line l in the image (Fig. 2) is given by*

$$u=\frac{n-(n,k)k}{\sqrt{1-(n,k)^2}}. \tag{2.2}$$

Lemma 3. *If n is the N-vector of a line l in the image, the unit vector v along the line l in the image (Fig. 2) is given by*

$$v=\frac{n\times k}{\sqrt{1-(n,k)^2}}. \tag{2.3}$$

Lemma 4. *If m is the N-vector of a point P in the image and n is the N-vector of a line l in the image, the distance $h(P,l)$ of point P from line l (Fig. 2) is given by*

$$h(P,l)=\frac{f}{\sqrt{1-(n,k)^2}}\frac{|(m,n)|}{(m,k)}. \tag{2.4}$$

Lemma 5. *If n is the N-vector of a line l in the image and m is the N-vector of a point P in the image, the N-vector m' of the point P' nearest to the point P on the line l (Fig. 2) is given by*

$$m'=N\left[k-(n,k)n+\frac{|m,n,k|}{(m,k)}n\times k\right], \tag{2.5}$$

where $|a,b,c|=(a\times b,c)=(b\times c,a)=(c\times a,b)$ denotes the scalar triple product of vectors a, b, and c.

3. Measure of Deviation for Primary Lines

Let us call the lines fitted to edges *primary lines*. By definition, a primary line has a finite part that was fitted to an edge. (Each edge segment defines one primary line. Fitting a common line to multiple edges is considered later.) Let us call it the *primary segment*. Let us assume that all edges are detected by the same procedure—application of the same edge operator, for example. Let us call the midpoint of the primary segment its *center*.

Suppose we expect a line $\bar{l}$ with N-vector $\bar{n}$ but instead observe a line l with N-vector n. (Throughout the rest of this paper, we use bars to denote quantities that are supposed to exist if no noise exists.) Let G be the center of line l, and w the length of the primary segment. Let $P(t)$ be a point on line l apart from the center G by distance t (the distance is signed appropriately), and let $m(t)$ be its N-vector.

Consider the integral

$$S=\int_{-w/2}^{w/2}h(P(t),\bar{l})^2dt, \tag{3.1}$$

where $h(P(t),\bar{l})$ is the distance of point $P(t)$ from line $\bar{l}$ defined by eqn (2.4) (Fig. 3). Let m_G be the N-vector of the center G of line l. From Lemmas 1, 3, and 4, it is easy to show that

$$S=\frac{w}{1-(\bar{n},k)^2}\left[f^2\frac{(m_G,\bar{n})^2}{(m_G,k)^2}+\frac{w^2}{12}\frac{|\bar{n},n,k|^2}{1-(n,k)^2}\right]. \tag{3.2}$$

Now, we introduce an approximation. Since l is a primary line, its primary segment actually appears in the original image. This means that line l passes near the image origin (compared with the focal length f). Since we are considering deviation of line l, line $\bar{l}$ is also expected to pass near the image origin. Hence, their N-vectors n and $\bar{n}$ are nearly parallel to the image plane. Thus, we have $(n,k)\approx 0$, and $(\bar{n},k)\approx 0$. Since both n and $\bar{n}$ are nearly parallel to the image plane, vector $\bar{n}\times n$ is nearly perpendicular to the image plane, and we have $|\bar{n},n,k|=(\bar{n}\times n,k)\approx\|\bar{n}\times n\|=\sqrt{1-(\bar{n},n)^2}$. Since the primary segment of line l actually appears in the original image, its center G is fairly close to the image origin (compared with the focal length f). Hence, its N-vector m_G is nearly perpendicular to the image plane, and we have $(m_G,k)\approx 1$.

Substituting these into eqn (3.2), we obtain

$$S\approx w\left[f^2(m_G,\bar{n})^2+\frac{w^2}{12}(1-(n,\bar{n})^2)\right]. \tag{3.3}$$

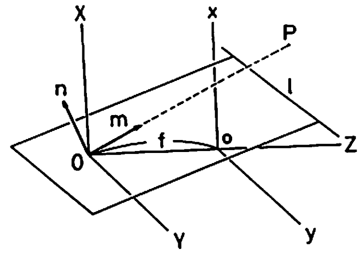

Fig. 1 N-vectors representing points and lines on the image plane.

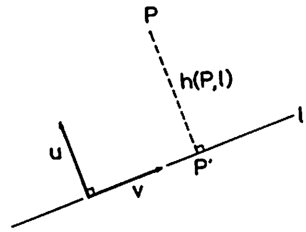

Fig. 2 Unit vector u is perpendicular to line l, and v is the unit vector along l. Point P' is the nearest point to point P on line l.

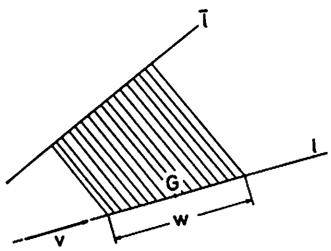

Fig. 3 The deviation of line l from line $\bar{l}$.

In view of this, we define the *measure of deviation* of line l from line $\bar{l}$ to be

$$D(l,\bar{l})=Cw\left[f^2(m_G,\bar{n})^2+\frac{w^2}{12}(1-(n,\bar{n})^2)\right], \tag{3.4}$$

where C is a factor reflecting the edge intensity and the edge width of the primary segment. Let us call it the *edge strength factor.*

4. Collinearity Test for Primary Lines

Given multiple primary lines, how can we decide whether or not these lines can be judged as identical? This problem is very important, because a straight object boundary is often detected as multiple fragmented edges due to the presence of noise. In the past, this problem of "edge grouping" has been treated in ad-hoc ways [1, 6].

Consider, for example, the two edges shown in Fig. 4(a). We can measure the distance δ and the angle α of deviation for each consecutive two edges, and judge the two edges as collinear if they are respectively smaller than appropriately set threshold values. Then, a different policy becomes necessary if we want to make a judgment about three or more edges (Fig. 4(b)).

Here, we take a consistent approach. Let us consider the case of Fig. 4(a). First, we *hypothesize* that the two line segments are collinear, and *estimate* a line $\bar{l}$ that supposedly contains the two edges. Then, we *test* this hypothesis by computing *to what extent these edges must be displaced if they all lie on the line* $\bar{l}$. The hypothesis is accepted if the deviation is smaller than a threshold value, and rejected otherwise.

The case of Fig. 4(b) can be treated in the same way; we first hypothesize a line that supposedly contains all the edges by the *same* fitting scheme, and then test this hypothesis by computing to what extent these edges must be displaced, using the *same* measure of deviation. This measure is thresholded by the *same* threshold value. Thus, we need not introduce any new criteria, measures, or threshold values. Although we must introduce the the measure of deviation and adjust the threshold value, they can be fixed for all types of problems.

A formal description is as follows. Let l_α, α=1, ..., N, be the primary lines to be tested for collinearity, and let n_α, α=1, ..., N, be their respective N-vectors. Let m_{G_α}, α=1, ..., N, be the N-vectors of their centers. The first step is hypothesizing a line $\bar{l}$. Here, we fit a line to the centers of the lines l_α weighted by $C_\alpha w_\alpha$ by an appropriate method (e.g., the least-square fitting), where w_α and C_α are the length and the edge strength factor of the primary segment of l_α. Let $\bar{n}$ be the N-vector of the fitted line.

The next step is testing this hypothesis. Let m_{G_α} be the N-vector of the center G_α of line l_α. Using eqn (3.6), we define the following *measure of collinearity*:

$$D(l_1,...,l_N;\bar{l})=\max_\alpha D(l_\alpha,\bar{l})$$

$$=\max_\alpha C_\alpha w_\alpha\left[f^2(m_{G_\alpha},\bar{n})^2+\frac{(w_\alpha)^2}{12}(1-(n_\alpha,\bar{n})^2)\right]. \tag{4.1}$$

The hypothesis is accepted if this value is below a fixed threshold value, and rejected otherwise.

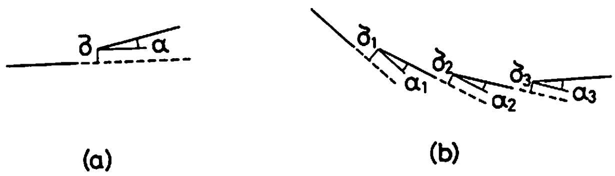

Fig. 4 (a) Two edges are judged as collinear if the angle α and the distance δ are separately smaller than appropriately set threshold values. (b) A different policy becomes necessary if we want to make a judgment about three or more edges.

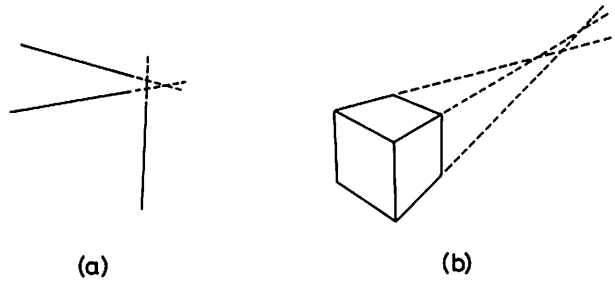

Fig. 5 (a) Edges that are supposed to meet at one corner vertex rarely define a single corner position. (b) Images of parallel line segments in the scene should meet, when extended, at a single vanishing point, but this cannot be expected for real images.

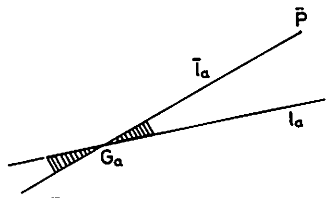

Fig. 6 Line $\bar{l}_\alpha$ passes through point $\bar{P}$ and the center G_α of line l_α.

5. Concurrency Test for Primary Lines

Consider how to judge whether or not multiple primary lines are concurrent. This problem frequently arises. For example, if we detect edges of an image of an object having corners, multiple (typically three) edges should precisely meet at one corner vertex. This, however, rarely happens for real images (Fig. 5(a)). Hence, we must make a judgment whether or not the given line segments define a single corner. Similarly, projections of edges parallel in the scene are concurrent, meeting, if extended, at a single *vanishing point.* However, we cannot expect this for real images (Fig. 5(b)). Hence, we must make a judgment as to whether or not the given line segments define a single vanishing point.

A naive way to solve this problem is to judge that lines are concurrent if the maximum separation of their intersections is below an appropriately set threshold value [2, 3]. However, the threshold value cannot be fixed. For example, tolerable errors must be severely restricted for the case of Fig. 5(a), while larger errors must be tolerated for the case of Fig. 5(b) because slight displacements of individual edges may cause a large deviation of their vanishing points.

Let l_α, α=1, ..., N, be the lines to be tested for concurrency, and let n_α, α=1, ..., N, be their N-vectors. First, we hypothesize that lines l_α all meet at a point $\bar{P}$ and estimate its

N-vector $\bar{m}$ by an appropriate method, e.g., the least-square method with each line l_α weighted by $C_\alpha w_\alpha$, where w_α and C_α are the length and the edge strength factor of the primary segment of l_α.

The next step is to test this hypothesis. We draw a line $\bar{l}_\alpha$ passing through $\bar{P}$ and the center G_α of line l_α (Fig. 6). It is easy to show that the measure of deviation, $D(l_\alpha, \bar{l}_\alpha)$, of line l_α from line $\bar{l}_\alpha$ is given by

$$D(l_\alpha, \bar{l}_\alpha) = \frac{C_\alpha}{12}(w_\alpha)^3\left[1 - \frac{|\bar{m} n_\alpha m_{G_\alpha}|^2}{1-(\bar{m}, m_{L_\alpha})^2}\right], \tag{5.1}$$

where m_{G_α} is the N-vector of the center G_α of line l_α.

Thus, we can test the concurrency by defining the following *measure of concurrency*:

$$D(l_1, ..., l_N; \bar{P}) = \max_\alpha D(l_\alpha, \bar{l}_\alpha)$$

$$= \max_\alpha \frac{C_\alpha}{12}(w_\alpha)^3\left[1 - \frac{|\bar{m} n_\alpha m_{G_\alpha}|^2}{1-(\bar{m}, m_{G_\alpha})^2}\right]. \tag{5.2}$$

The hypothesis is accepted if this value is below a fixed threshold value.

6. Collinearity Test for Secondary Points

Consider how to make a judgment as to whether or not a given set of points is collinear. This problem occurs in many problems. For example, if horizontally placed object faces exist in the scene and they have parallel boundary edges, their vanishing points must all lie on a common *horizon* (Fig. 7). Hence, we can check whether individual faces are horizontal or not by testing collinearity.

However, exact concurrency cannot be expected for real data. Moreover, since points are computed data, they can be located *anywhere on the image plane*. The judgment must be done in a *finite* domain of computation, and the tolerance must depend on the positions and the configuration of these points. These issues are solved by the use of N-vectors and our strategy of hypothesizing and testing.

Let P_α, $\alpha=1, ..., N$, be the points to be tested, and let m_α, $\alpha=1, ..., N$, be their N-vectors. The first step is to hypothesize a line $\bar{l}$ that supposedly passes through these points by an appropriate (e.g., the least-square) method. The next step is testing this hypothesis. Recall that each point is defined as the intersection of some lines. Let $l_\beta^{(\alpha)}$, $\beta=1, ..., N_\alpha$, be the primary lines that define point P_α, and let $n_\beta^{(\alpha)}$, $\beta=1, ..., N_\alpha$, be their N-vectors. In other words, point P_α is given as the common intersection of primary lines $l_\beta^{(\alpha)}$, $\beta=1, ..., N_\alpha$.

First, let $\bar{P}_\alpha$ be the point on the hypothesized line $\bar{l}$ that are closest to point P_α. The N-vector $\bar{m}_\alpha$ of point $\bar{P}_\alpha$ is computed by eqn (2.5) of Lemma 5. Let $w_\beta^{(\alpha)}$ and $C_\beta^{(\alpha)}$ be the length and the edge strength factor of the primary segment of line $l_\beta^{(\alpha)}$, and $m_{G_\beta^{(\alpha)}}$ the N-vector of its center $G_\beta^{(\alpha)}$. We use the following *measure of collinearity* (Fig. 8):

$$D(P_1, ..., P_N; \bar{l}) = \max_\alpha D(l_1^{(\alpha)}, ..., l_{N_\alpha}^{(\alpha)}; \bar{P}_\alpha)$$

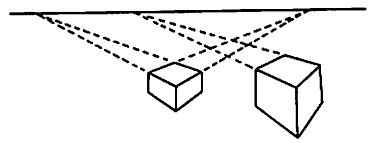

Fig. 7 The vanishing points of parallel edges of horizontally placed objects must all lie on a common *horizon*.

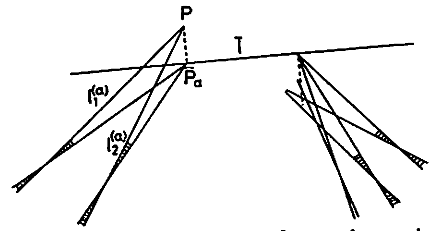

Fig. 8 The measure of collinearity of secondary points P_1, ..., P_N.

$$= \max_\alpha \max_\beta \frac{C_\alpha}{12}(w_\beta^{(\alpha)})^3\left[1 - \frac{|\bar{m}_\alpha n_\beta^{(\alpha)} m_{G_\beta^{(\alpha)}}|^2}{1-(\bar{m}_\alpha, m_{G_\beta^{(\alpha)}})^2}\right]. \tag{6.1}$$

Thus, we are measuring *to what extent the original edges must be displaced* so that the intersections all lie on the hypothesized line $\bar{l}$. The hypothesis is accepted if this measure is below a fixed threshold, and rejected otherwise.

7. Concluding Remarks

In this paper, we have presented a general formalism for testing geometric configurations of image data by taking into account the *hierarchy* of image data resulting from image processing procedures. The basic principle is *hypothesizing* and *testing*: We first estimate an ideal geometric configuration that supposedly exists, and then check *to what extent the original edge data must be displaced* in order to support the hypothesis. All types of tests are reduced to computing a single *measure of edge displacement*, which provides a universal *measure of uncertainty* applicable to all types of decision-making. In our formalism, no explicit forms of probability distribution need be introduced, and all the procedures are described by explicit algebraic expressions in N-vectors.

References

[1] K. M. Andress and A. C. Kak, Evidence accumulation & flow of control in a hierarchical spatial reasoning system, *AI Magazine*, 9(2) (1988), 75-94.

[2] K. Kanatani, Reconstruction of consistent shape from inconsistent data: Optimization of 2½D sketches, *Int. J. Comput. Vision*, 3(4) (1989), 261-292.

[3] K. Kanatani, *Group-Theoretical Methods in Image Understanding*, Springer, Berlin 1990.

[4] M. J. Magee and J. K. Aggarwal, Determining vanishing points from perspective images, *Comput. Vision Graphics Image Process*, 26 (1984), 256-267.

[5] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Vols. 1, 2, 2nd ed., Academic Press, Orlando, FL, 1982.

[6] R. Weiss and M. Boldt, Geometric grouping applied to straight lines, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, June 1986, Miami Beach, FL, pp. 656-661.