[5] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
[6] H. Derin and W. S. Cole, "Segmentation of textured images using Gibbs random fields," *Comput. Vision, Graphics, Image Processing*, vol. 35, pp. 72–98, 1986.
[7] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 1, pp. 39–55, Jan. 1987.
[8] C. S. Won and H. Derin, "Unsupervised segmentation of noisy and textured images using Markov random fields," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 4, pp. 308–328, July 1992.
[9] F. S. Cohen and Z. Fan, "Maximum likelihood unsupervised textured image segmentation," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 3, pp. 289–251, May 1992.
[10] N. Ahuja and A. Rosenfeld, "Mosaic models for textures," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, no. 1, pp. 1–11, Jan. 1981.
[11] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, no. 1, pp. 25–39, Jan. 1983.
[12] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Statist. Soc., Series B*, vol. 36, pp. 192–236, 1974.
[13] ——, "On the statistical analysis of dirty pictures," *J. Roy. Statist. Soc., Series B*, vol. 48, pp. 259–302, 1986.
[14] C. O. Acuna, "Texture modeling using Gibbs distributions," *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 3, pp. 210–222, May 1992.

Contribution to the Determination of Vanishing Points Using Hough Transform

Evelyne Lutton, Henri Maître, and Jaime Lopez-Krahe

Abstract—**We propose a method to locate three vanishing points on an image, corresponding to three orthogonal directions of the scene. This method is based on two cascaded Hough transforms. We show that, even in the case of synthetic images of high quality, a naive approach may fail, essentially because of the errors due to the limitation of the image size. We take into account these errors as well as errors due to detection inaccuracy of the image segments, and provide a method efficient, even in the case of real complex scenes.**

Index Terms—**Bias and errors of the Hough transform, Hough transform, orthogonal directions detection, vanishing points detection.**

I. Introduction

In many tasks of artificial vision, an accurate location of vanishing points is a first step toward three-dimensional (3D) interpretation. Vanishing points are defined in the image plane as those points where the images of all 3D scene lines, parallel to some space direction, converge. To one 3D space direction is attached one vanishing point on the image plane and conversely.

Detection of vanishing points, which is of little help in natural outdoor scenes, becomes of prime importance in the man-made environment where regular block looking structures or parallel alignments (streets, pavements, railroad) abound. We have the Italian Quattrocento to thank for the deep comprehension of the formation of perspective images,[1] a comprehension that has been a constant preoccupation of the theoricians in aesthetics up to the 20th century.[2] Traces of rigorous mathematical bases are found mainly at the corner of the 18th–19th centuries.[3, 4] Although the geometrical construction of vanishing points may become very complex when no hypothesis is made on the vision system, in the case of a perfect conic projection (pinhole cameras) it may be solved easily, since the image of a straight line remains a straight line. We will stay in the assumption of conic projection throughout this paper.

Manuscript received October 9, 1990; revised September 17, 1992. This work was supported by DRET and in collaboration with MATRA.

The authors are with Departement Images, Telecom Paris, Cedex, France.

IEEE Log Number 9206561.

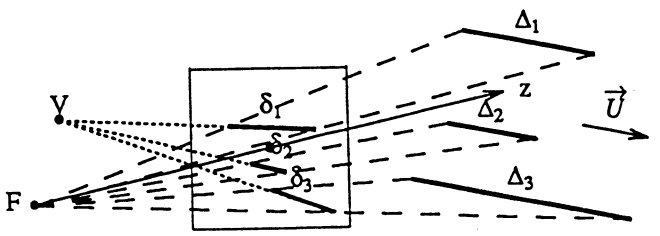

Fig. 1. Δ_1, Δ_2, and Δ_3 are lines of the 3D space parallel to the direction $\vec{U}$. δ_1, δ_2, and δ_3 their images on the image plane. δ_1, δ_2, and δ_3 converge to the vanishing point associated to $\vec{U}$.

Most of the existing methods to detect automatically these vanishing points stand on the use of the Hough transform, explicitly or not [16], [20], [21]. Hough transform is a global technique for detecting parametrical structures in images [5], [7], [17]. Some primitives are detected in the image, and then mapped into a parameter space; underlying structures are detected by searching for clusters in this parameter space. Two methods can be distinguished to fill the accumulators of Hough space [17].

- The "one-to-many" (1-to-m) transform, used most of the time, where for each feature point in the picture plane, several accumulator cells are incremented.
- The "many-to-one" (m-to-1), where we make use of several feature points in the image plane to increment exactly one accumulator cell.

For the application of vanishing points detection, the primitives are line segments. The vanishing points are thus characterized as those points where most of the supporting lines of these segments intersect (Fig. 1). Most of the time, these points are located far away from the image limits and even can be at infinity (for frontal lines).

So, the most important problem of the detection of vanishing points with the help of the Hough transform is the choice of the Hough space parameterization. Two main orientations have appeared in the literature, following the 1-to- m or the m-to-1 transforms; they are chronologically:

The 1-to-m approach with:

- Kender [9] in 1979 who uses, directly on the image plane, circles passing through the origin, and proposes either a search in a tridimensional space or two successive transforms;
- Ballard and Brown [2] in 1982, who propose a $(k/r, \theta)$ parameterization, with k constant, for the image primitives, which permits restriction of parameter space;

[1] L. B. Alberti, "De pictura" (manuscript 1435), printed in Basl (Swiss), 1540.

[2] E. Panovsky, "Die Perspektive als symbolische Form," Berlin, 1927.

[3] G. Monge, "Géométrie descriptive," *Leçons données aux Ecoles Normales de l'an III de la République*, Paris, 1798.

[4] J. V. Poncelet, "Traîté des propriétés projectives des figures," Paris, 1822.

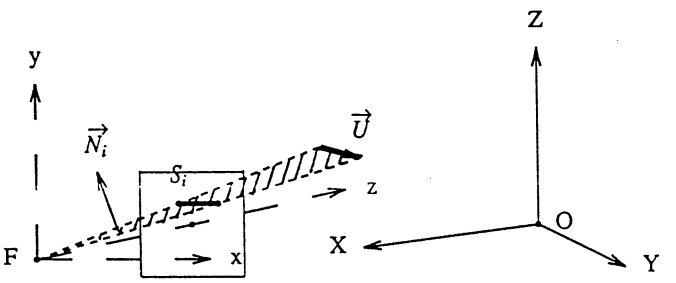

Fig. 2. Construction of the image: $\vec{N}_i$ is the normal vector of the interpretation plane of segment S_i. (O, X, Y, Z) is the coordinate frame of the 3D scene, (F, x, y, z) of the camera.

- Barnard [3] in 1983, who first proposed the use of the interpretation planes of the image segments, and the use of the Gaussian sphere centered on the optical center of the camera as parameter space. This formulation has the great advantages to represent as well the finite and infinite vanishing points, to furnish a restricted Hough space; and
- Quan and Mohr [20] in 1989, who propose an improvement of the Barnard method using a hierarchical method to divide the Hough space.

On the other side, equivalent to the m-to-1 approach, the method of Badler [1] in 1974 and, later, of Magee and Aggarwal [16], in 1984, which are based on cross products of normal vectors to interpretation planes, and uses the (θ, φ) parameterization of the Gaussian

Let us notice, that these two families of methods converge into the use of the Gaussian sphere as parameter space.

Our concern is to detect vanishing points in man-made environments where parallel lines abound. More specifically, we are interested in detecting the three main orthogonal directions that are often used as the frame for the construction of human productions: for instance, it is the case of some outdoor scenes (streets in urban environment, factories, dockyard, airports in aerial images), of some indoor scenes (rooms and corridors, offices, and workshops), and even of consumer products. (Examples of images with such dominant directions are presented in Figs. 13–15.)

The method we propose is in two steps (Section II); the first ignores the three-orthogonal-direction hypothesis, and is similar to [3]. The second exploits the three-orthogonal-direction hypothesis by means of a second Hough transform. When using the method in [3], we underline the intrinsic limitations due mostly to the finite extension of the image. We show that these limitations may prevent the extraction of the main directions even in the case of simple objects without noise. By theoretic analysis of the role of the image dimension (Section III), we conclude by proposing a more robust algorithm. Results are presented on several different images supporting our demonstration (Section IV). A discussion and a comparison with the method [20] are given in Section V.

II. Principles of the Method

A. Step 1: Direction Determination

1) Using the Gaussian Sphere: When searching for vanishing points, instead of looking for the convergence of lines on the image, we can look for the intersection of circles on the Gaussian sphere associated to the image [3]. We denote by (O, X, Y, Z) the coordinate frame associated with the 3D scene, and by (F, x, y, z) the coordinate frame of the image, centered on the optical center of the camera (see Fig. 2).

For each line of the image plane that supports a segment S_i of the image, a great circle can be drawn on the Gaussian sphere Γ, centered on the optical center of the camera, F. This circle represents all the

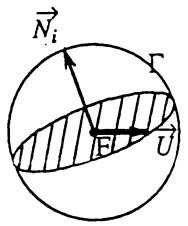

Fig. 3. Gaussian sphere: The great circle represents the interpretation plane of S_i on Γ. All the $\vec{U}$ vectors have an image on the supporting line of S_i.

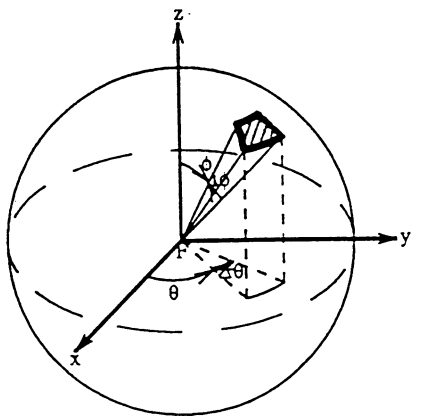

Fig. 4. Rectangular cell of size $\Delta\varphi, \Delta\theta$.

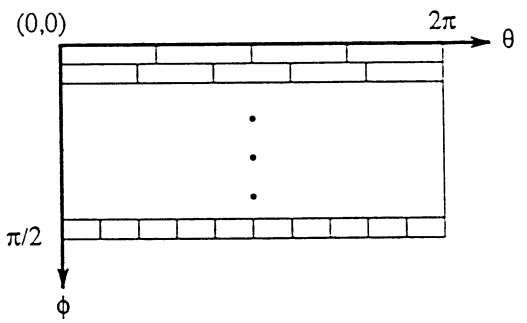

Fig. 5. Almost-equal-size cells quantization.

3-D directions $\vec{U}$ that project on the image line. It is the trace of the interpretation plane of the line (or of the corresponding segment) on the Gaussian sphere (Figs. 2 and 3). A direction $\vec{U}$ that can be projected on S_i is $\vec{N}_i \cdot \vec{U} = 0$.

Using the spherical coordinates, for a known vector $\vec{N}_i(\theta_i, \varphi_i)$, all the possible vectors $\vec{U}(\theta_u, \varphi_u)$, verify the equation of a great circle on Γ: $\cos(\theta_i - \theta_u)\sin\varphi_i \sin\varphi_u + \cos\varphi_i \cos\varphi_u = 0$.

To every line intersection on the image plane corresponds one circle intersection on the Gaussian sphere, which represents a 3D direction of the scene, or one vanishing point. With this representation, problems of infinite vanishing points are avoided, since Γ is a closed representation of R^2. As the problem is symmetrical with respect to the (x, y) plane, only one-half of the Gaussian sphere is used: $\theta_u \in [0, 2\pi[, \varphi_u \in [0, \pi/2[$.

2) Quantization of Hough Space: As Barnard [3] noticed, a regular quantization in θ and φ does not provide cells of equal surfaces on Γ. Therefore, a regular quantization will not allow a straightforward determination of the maximal direction probability from the maximal cell count. A postprocessing of cell counts is needed to determine probabilities. Another solution is to use an equal-size-cell quantization, a solution reputed not flexible on the sphere since it uses triangle-shaped cells and constrains the cell size drastically.

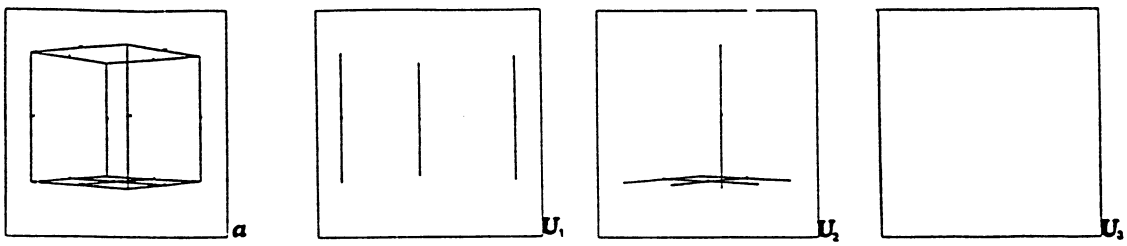

Fig. 6. Failure on a simple artificial case: a, synthetized image segments; $\vec{U}_1$, detected vertical parallel to the image plane; $\vec{U}_2$, first horizontal direction, pointing into the image (all the segments cross each other in a small area near the center of the image); and $\vec{U}_3$, second horizontal direction (no associated segments).

For these reasons, we propose a quantization in "semiregular" rectangular cells, based on a regular quantization in ϕ and an irregular quantization in θ. On the Gaussian sphere, a cell of size $\Delta\phi, \Delta\theta$ has a surface of $\Delta S = \Delta\theta[\cos\rho - \cos(\rho + \Delta\rho)]$, see Fig. 4. Let N_T be the number of layers between 0 and $\pi/2$ for $\rho : \Delta\rho = \pi/(2N_T)$.

For a chosen ΔS (fixed by the ultimate accuracy we want), we construct cells on the sphere corresponding to the kth layer, with a spacing along θ equal to

$$\Delta\theta(k) = \frac{\Delta S}{\Delta\rho} \cdot \frac{1}{\cos(k-1) - \cos(k)}.$$

Such an accumulator on the (θ, ρ) plane has almost-equal-size cells. Therefore, if the directions $\vec{U}$ were in infinite number and distributed uniformly, a uniform count on each cell would be obtained. Fig. 5 gives an image of this quantization. Thus, the search of accumulation points on the Gaussian sphere is reduced to the search of local maxima on this "almost regular" accumulator.

B. Step 2: Three-Orthogonal-Directions Detection

1) Searching for a Vertical Direction: After the previous step, we have a list of directions $\vec{U}_k$, each of them corresponding to a possible vanishing point in the image plane and ranked by their counts in the accumulator space. Now we want to select, from this set, the most represented three orthogonal directions. At this point, we make use of a supplementary assumption that is dictated by the observation of man-made environments. Not only in man-made environments three orthogonal directions are often dominant, but also, one of them is orthogonal to most of the directions of the scene. This direction will be referred to as vertical[5] (and denoted as $\vec{V}$) since it is generally the case.

Thus, we search for a direction $\vec{V}$, orthogonal to the greatest number of $\vec{U}_k$ directions. This stage is done by a second Hough transform, identical to the first one. For each direction $\vec{U}_k$ is drawn the orthogonal great circle on a second Gaussian sphere Γ'. The vertical $\vec{V}$ is with high probability at the most represented intersection point on Γ'. This method provides the vertical direction even when this direction is not represented by any segment in the set of $\vec{U}_k$.

The next step, to detect the two other directions, is then straightforward. The direction $\vec{V}$ provides the first requested direction $\vec{U}_1(\vec{V} \equiv \vec{U}_1)$. Then, in the set of $\vec{U}_k$ orthogonal to $\vec{U}_1$, the two most represented orthogonal directions $\vec{U}_2$ and $\vec{U}_3$ are selected.

C. Conclusion

The principles of the detection method presented here appear sound. Unfortunately, when used carelessly, the detection does not work correctly. Failures exist not only on complex real scenes when noise strongly disturbs the signal, but even on simple artificial school cases. Fig. 6 shows such an example, where the most important vanishing direction $\vec{U}_2$ has been interpreted as pointing inside the image, the vertical $\vec{U}_1$ is parallel to the image plane, and even, the third direction $\vec{U}_3$ does not correspond to any segment. We will see in the next section that this is mostly due to the limitation of the image.

[5] *Remark:* We see from the previous lines, the "vertical" is given, in this method, as the direction which is the most orthogonal to the other lines; it does not deal with gravity, and we are not claiming to detect the direction along which apples fall, unless the scene is favorable!

III. Study of Errors

To understand these problems, it is necessary to look carefully at the possible errors and biases introduced during the Hough transform. They have three different origins [18]:

- bias due to the finite extension of the image (the image size is limited) [4], [18];
- errors due to the quantization of the images [8], [12], [13], [24];
- errors due to the detection inaccuracy of the image primitives [6], [18], [22], [23].

In case of vanishing point detection, we have noticed experimentally that the most important is the first error, which favors the apparition of artifacts. As a result of the finite extension of the image, the probability to detect a vanishing direction anywhere on the Gaussian sphere is nonuniform, and the vanishing directions that give vanishing points inside the image are favored. The second error source is negligible as long as large enough images are taken. The third one may receive easy solutions (see Section III-B).

A. Bias Due to Image Limitation

For simplicity, we suppose that the image is circular. A great circle drawn on the Gaussian sphere is interpreted as the projection on the Gaussian sphere of the line supporting the image segment. However, only those lines seen in the image create a circle; thus, only circles that pass through the projection of the image on Γ (dotted spherical cap of Fig. 7) are drawn on Γ. This means that with high probability intersections of circles will be found in this cap, corresponding to vanishing points inside the image itself. Fortunately this problem may be solved analytically.

1) A Priori Probability of Direction $\vec{U}$: To compensate for a bias to the positioning of the vanishing point in the image, we use the hypothesis that we have a uniform distribution of 3D segments in the scene; that is, every vanishing point has an equal probability to be found in the image. An exact calculation of the a priori probability of a direction $\vec{U}$ in the Hough accumulator has been derived, in the case of a circular retina (see Appendix I). In the case of a circular image with projection angle ρ_0 on Γ, this probability being symmetrical with respect to the visual axis z can be expressed as a function of the angle ρ, the angle between the direction $\vec{U}$, and the z axis (see Fig. 8). The detection of a priori probability density of a direction $\vec{U}$ is given by

$$\begin{aligned} P(\vec{U}) &= \frac{1}{S}, && \text{if } \rho_U \le \frac{\pi}{2} - \rho_0, \\ &\text{with } S = \int_0^{\pi/2} P(\vec{U})\, d\vec{U} \\ P(\vec{U}) &= \frac{2}{\pi S} \arcsin\left[\frac{\cos\rho_0}{\sin\rho_U}\right], && \text{if } \rho_U \ge \frac{\pi}{2} - \rho_0. \end{aligned} \tag{1}$$

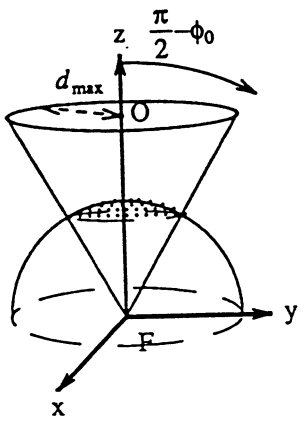

Fig. 7. Projection of a circular image on Gaussian sphere.

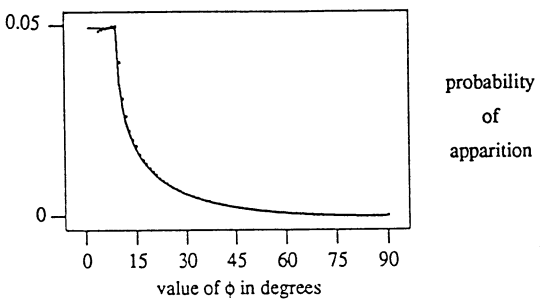

Fig. 8. Probability curves (the simulated curve id is dotted and the theoretical is continuous) for a circular image $d_{max} = 512$ pixels, with a pixel size of $25 \times 25\,\mu m$ and a focal length of 40 mm.

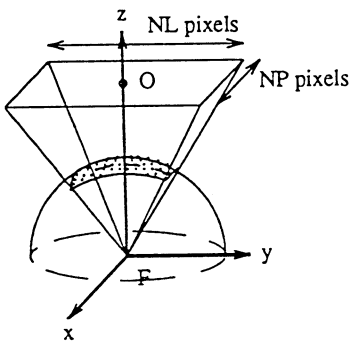

Fig. 9. Projection of a rectangular image on Gaussian sphere.

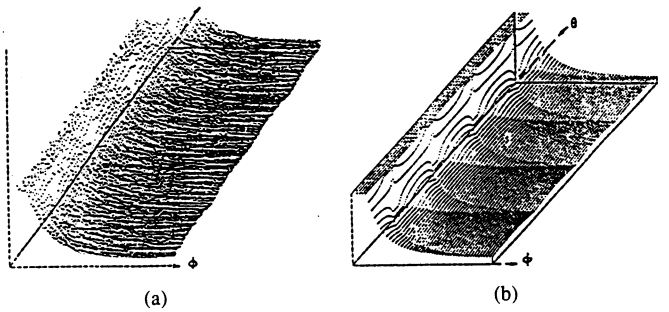

Fig. 10. Probability surfaces for a square image 512×512 pixels, with pixel size of $25 \times 25\,\mu m$ and a focal length of 40 mm. (a) Simulated surface. (b) Theoretical surface.

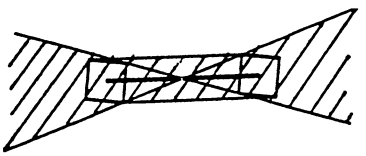

Fig. 11. Error on the direction of a segment.

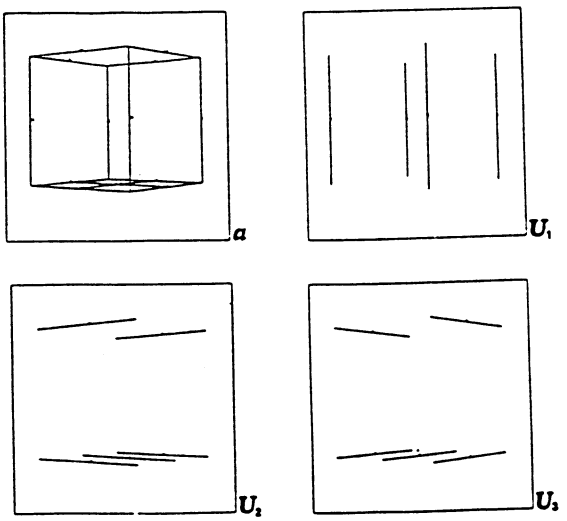

Fig. 12. Simple artificial case: The cube a, synthesized image segments; U_1, detected vertical; U_2 and U_3, the two horizontal directions (see Fig. 6).

In the case of a rectangular retina, the probability is no longer symmetrical with respect to the z axis; it has a complex expression, a function of the position of the line with respect to the borders (see Appendix I).

The surface of probability density is shown in Fig. 10 as a function of the spherical coordinates of the Gaussian sphere θ and ρ. We notice that, here again, the projections inside the retina are favored on the Gaussian sphere, but now, outside the projection of the image, the directions $\theta = k(\pi/2)$ are slightly favored.

With these tools we may compute the impulse response of the Hough transform on the first accumulator. It should be spatially invariant and look like a Dirac function. This is true in case of an infinite retina, but no longer for the case of a finite-extent retina (see Appendix III). A similar estimation of the bias has to be made for the second Hough transform, whem computing the a priori probability of a vertical $\vec{V}$ (see Appendix II). The limitation of the retina has the effect that the verticals parallel to the image plane are also favored.

2) Verifications with Simulation: These theoretical results are confirmed by simulations. Simulated images have been built with segment projections on circular or rectangular images of Poisson-law random 3D lines. Simulated and computed curves and surfaces fit nicely (Figs. 8 and 10).

3) Correction: Using this probability it becomes possible to eliminate the undesirable effect of a restricted retina. Of course, a real 3D scene is not quite amenable to the previous discussion. A complex scene is made of segments belonging to the set of three orthogonal directions, and of segments that are just noise with respect to this signal. We will consider that this noise is distributed randomly, and, depending on its relative importance, two solutions are intuitively possible: We may either subtract a judiciously weighted a priori probability from the experimentally obtained accumulators or divide each cell by this probability.

Let $A(I, J)$ be the count of the cell (I, J), J corresponding to ρ and I to θ. Let $P(I, J)$ be the probability of apparition of a vanishing direction in the cell (I, J), computed from the theoretical formulas. We may create a new accumulator whose cells are either $A(I, J) - K \cdot P(I, J)$ or $A(I, J)/[K \cdot P(I, J)]$, where $K \cdot P(I, J)$ is an estimation of the accumulator if there were no vanishing points in the image (random behavior). The factor K is related to a signal-

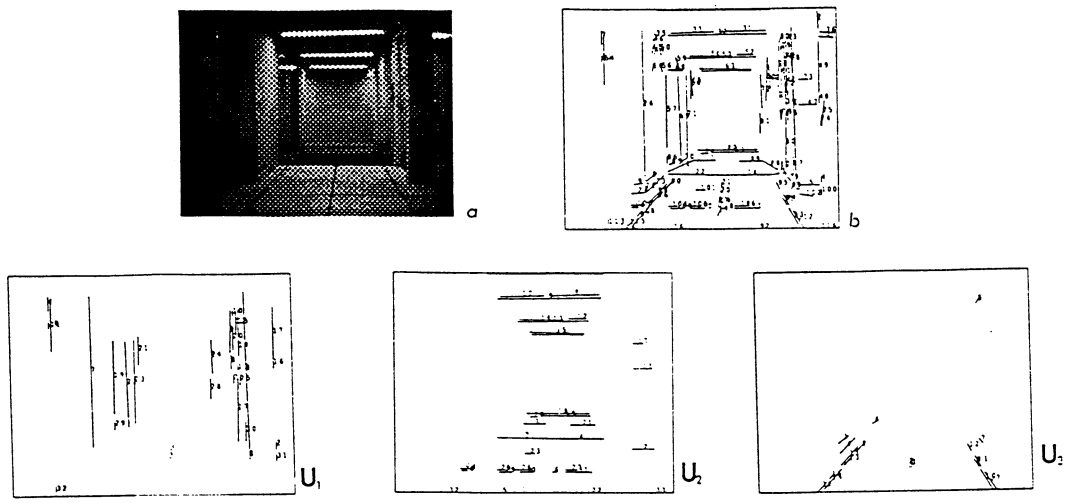

Fig. 13. (a) Original image (taken by a CCD camera, $f = 50$ mm, pixel size = $100 \times 100\,\mu$m). (b) Segments of the image: U_1, detected vertical; U_2 and U_3, the two horizontal directions (courtesy INRIA).

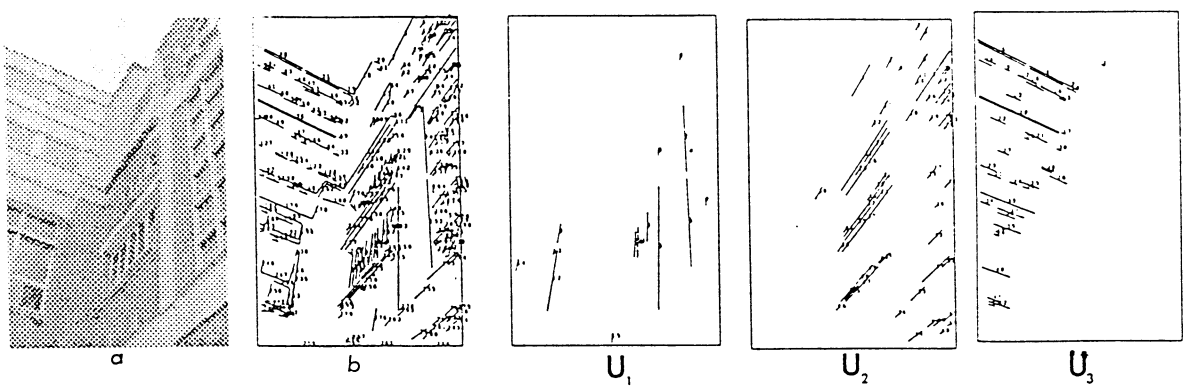

Fig. 14. (a) Original image ($f = 50$ mm, pixel size = $100 \times 100\,\mu$m). (b) Segments of the image: U_1, detected vertical; U_2 and U_3, the two horizontal directions.

to-noise ratio of the image segments; it is experimentally difficult to estimate. Fortunately, in the second expression it has no influence on the detection of the maximal counts and thus may be neglected. We have experimented with both solutions. The second solution, which does not need any estimation of the noise, has been verified experimentally as the most efficient. It has been used for the results presented below.

B. Errors on the Position of the Image Segments

The other noticeable source of errors is due to errors in the location of image primitives. This effect is smaller than the one due to the retina limitation, but influences on the accuracy of results. A quantization error on each extremity of a segment induces an error on the direction of the segment that can be very important, especially when the segment length is short (see Fig. 11).

Iannino [6] has proved, in the case of detection of lines and circles in an image, that there is a threshold, related to the errors on the primitives, under which it is useless to quantify the Hough space. Indeed, under this threshold, due to quantization errors on the image, the detection of local maxima of the accumulator is mistaken. In the problem of vanishing points detection, the calculus is not straightforward because of the nonhomogeneity of the errors on primitives. However, we can estimate easily an approximate mean value of this threshold [15].

As a matter of fact, we have adopted a strategy to update the cells of the accumulator that are not on a curve (representing the great circle and so the interpretation plane), but in a zone around this curve, representing all the interpretation planes of the possible real segments enclosed in the hatched area of Fig. 11. The weight, proportional to the length of the segment, is distributed uniformly on all the connected cells. This method allows one to take small segments into account, and to grant greater confidence in long ones. And, if we are near the Iannino threshold, we can assume that for the long segments we update the minimum number of cells.

IV. Experimental Results

First the result of our algorithm on the school example of Fig. 12 is shown: all the segments are now classified correctly.

The following examples show the results of the algorithm on real images. First, the results are shown in Fig. 13 of the algorithm run on an indoor image, taken with a CCD camera. Edges are detected using a Sobel detector, then thinned, and fitted to segments. The segments of length less than 10 pixels are rejected, 116 segments are kept for vanishing points detection, 83 of them are found as participating to the main directions.

Fig. 14 is a photograph of buildings, taken with a conventional camera on film and digitized. Two hundred eighty-six segments were detected on the image and 118 recognized as converging into the three main vanishing points.

Finally, Fig. 15 is an aerial photograph of industrial buildings; 183 segments, from 266 detected, are recognized as converging to one of the three vanishing points.

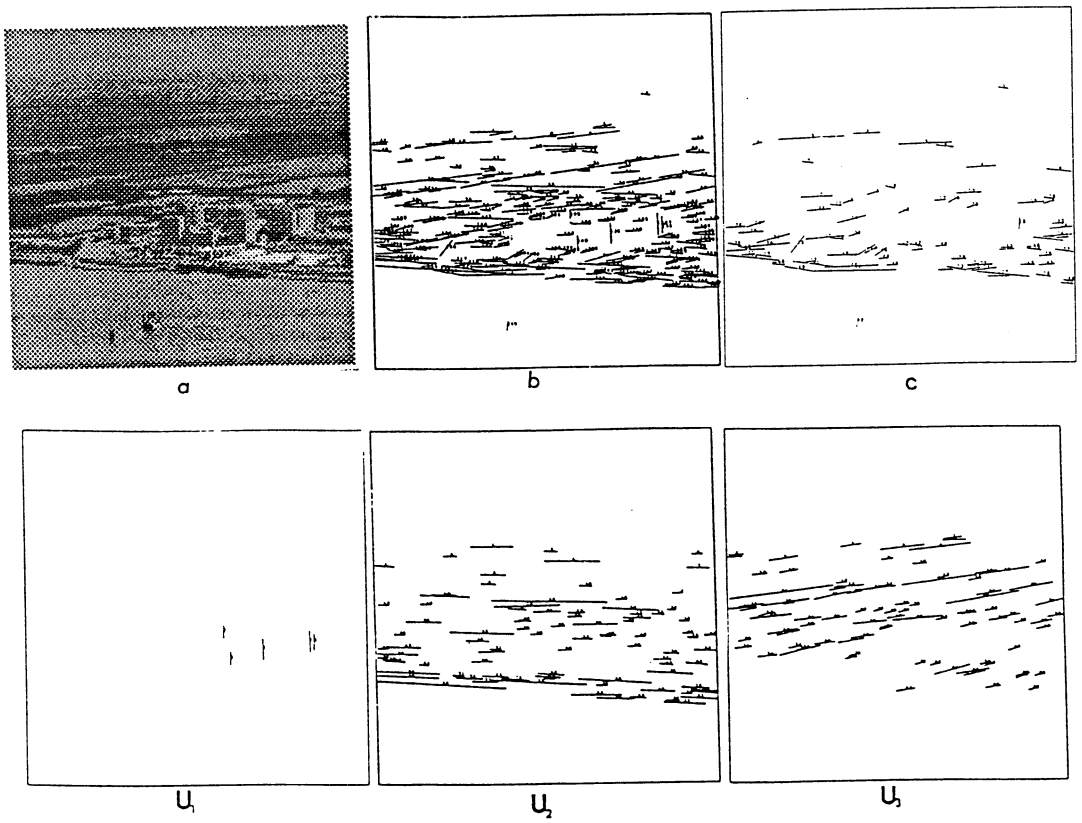

Fig. 15. (a) Original image ($f = 100$ mm, pixel size = 25 μm). (b) Segments of the image: U_1, detected vertical; U_2 and U_3, the two horizontal directions. (c) Not classified image segments (courtesy MATRA-MS21).

V. Discussion

A. Searching for the Associated Segments

For each of the three directions $\vec{U}_k, k \in 1,2,3$, the associated image segments are the segments whose normal vectors to the interpretation plane $\vec{N}_i$ verify $\vec{N}_i \cdot \vec{U}_k = 0$. In fact, due to inaccuracy all along the processing, we cannot test the equality; we use a threshold ε: $|\vec{N}_i \cdot \vec{U}_k| \leq \varepsilon$. It seems reasonable that the threshold varies with respect to the distance between the segment and the vanishing point in the image plane. Indeed, for example, when the vanishing point is in the image, the projections of the real scene segments appear as very short (and noisy) segments around the vanishing point. So if ε is too large, all the image segments around the vanishing point verify the criterion (because all the directions in the image plane are allowed near the vanishing point), and a lot of false segments are associated. For this reason, we take an ε small when the distance between the segment and the vanishing point is small, and larger when this distance increases.

B. Calibration of Camera Internal Parameters

The efficiency of the accumulator correction has been shown in the previous examples, but this efficiency relies on an exact knowledge of the geometrical setup. More specifically, we need to know the exact focal length and the center of view to run our algorithm correctly. If the focal length is mistaken, the results of the search of three orthogonal scene directions are false, because

1) The three vanishing points have associated vanishing directions on the Gaussian sphere that are not orthogonal to each other (see Fig. 16), and
2) The retina projection on the Gaussian sphere is changed, so that the correction function is altered and does not correct the accumulator in the right way.

For the same reasons, we need to know precisely where the center of view lies, that is, the intersection of the optical axis with the image plane. If the center of view is not correctly known, the correction function is skewed, and the angles are also misinterpreted. It must be remarked that the lack of precision on the center of view is especially sensible when vanishing points are in the image or near the limits of the image. Indeed such vanishing points need very strong criteria to search for the associated segments.

C. Comparison with Quan and Mohr Method

Quan and Mohr [19], [20] proposed a method to detect the vanishing points, which is an extension of the Barnard method [3]. This method works with the help of an accumulator defined on the Gaussian sphere according to Barnard's principles, but with a different quantization. Quan and Mohr build the accumulator by hierarchical division (Li and Lavin method [10], [11]) of the half Gaussian sphere according to the spherical parameterization (θ, ϕ). They do not mention any correction of bias. The experimentations of this algorithm[6] permit us to understand its behavior with bias due to the finite extension of the retina. It is particularly related to the minimal resolution R_0 used to limit the division of accumulator cells. In fact, if a minimal resolution R_0 near the one we use in our algorithm (about 1°) is fixed, the false detections due to the limitation of the retina appear, whereas if the minimal resolution R_0 is much smaller, the most represented vanishing points are detected correctly. This is because the searched local maxima are very high and fine peaks so that the bias due to the limitation of the retina does not alter the detection. Those peaks correspond to the impulse response we have calculated (see Appendix III).

[6] We would like to thank Audigier and Trullen, who programmed the Quan and Mohr method.

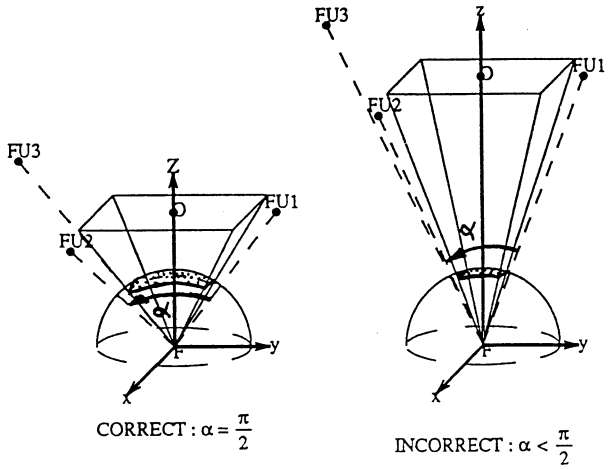

Fig. 16. Influence of focal length.

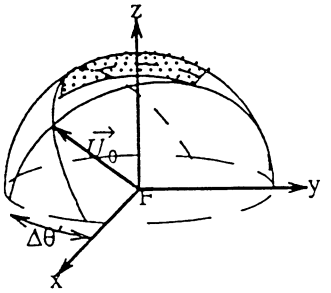

Fig. A1. Allowed zone for great circles.

Yet, in the case of a noisy image, that is, when the segments are relatively short, the "blurring" effect on the maxima, due to the quantization errors, extends the width of the peaks. Thus artifacts may appear. It is the case of aerial images (Fig. 15) on which we do not obtain good results with the Quan and Mohr algorithm.

The Quan and Mohr algorithm needs at least three passes and does not guarantee the orthogonality of the recovered vanishing directions. Note that, as a result of these three passes, it is not crucial for them to have the focal length f: however, Quan and Mohr argue [20] that the choice of f has an influence on the detection accuracy and that the exact focal length f is needed to estimate the real 3D direction. However, their algorithm seems to be less sensitive than ours to the calibration precision of f.

VI. Conclusion

We have proposed a complete method to improve the detection of vanishing points, and to detect three orthogonal vanishing directions even if one is not or nearly not represented among the image segments. This method works on images of real scenes, where many segments, some converging to vanishing points and some not, are detected. It provides excellent results on cases where methods were failing that do not correct the bias due to limitation of the retina.

Appendix I
A Priori Probability Density of Direction $\vec{U}$

A. Circular Image

First, the calculation is made in the case of a circular image (of radius $d_{\max}$): that is, where all the image lines verifying $d \leq d_{\max}, \forall \theta$, are allowed (Fig. 7). $o_0 = \arctg(f/d_{\max})$ denotes the maximal angle with z axis of interpretation plane normal of segments of this image.

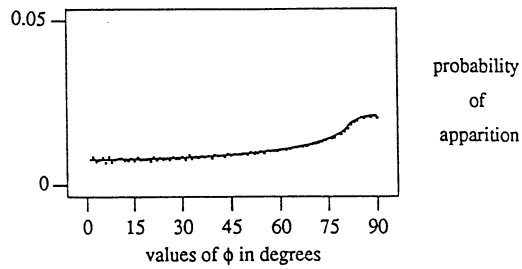

Fig. A2. Probability curves (simulated curve is dotted) for a circular image $d_{\max} = 512$ pixels, with a pixel size of $25 \times 25\ \mu$m and a focal length of 40 mm.

Suppose each direction of the Gaussian sphere is represented continuously and uniformly in the image. The probability of detection of a direction $\vec{U}$ is related to the "number" of great circles that pass through the point representing $\vec{U}$ on the Gaussian sphere. This probability is proportional to the surface containing all the possible great circles passing through the projection of the retina and through the point $\vec{U}$.

This surface is either the total Gaussian sphere surface if $o_{U} \leq (\pi/2) - o_0$, or, if not, a part of the sphere surface located between two planes passing through $\vec{U}$ and supported by the projection of the retina (Fig. A1). These two planes make an angle $\Delta\theta$.

Therefore, the probability is proportional to the ratio of this surface by the surface of the Gaussian sphere: if $o_{U} \leq (\pi/2) - o_0$, $S(\vec{U}) = 1$: if $o_{U} \geq (\pi/2) - o_0$, $S(\vec{U}) = (\Delta\theta/2\pi)$. We obtain

$$S(\vec{U}) = \begin{cases} 1. & \text{if } o_{U} \leq \dfrac{\pi}{2} - o_0 \\ \dfrac{2}{\pi} \arcsin\left[\dfrac{\cos o_0}{\sin o_{U}}\right]. & \text{if } o_{U} \geq \dfrac{\pi}{2} - o_0 \end{cases}$$

where

$$D(\vec{U}) = \frac{S(\vec{U})}{\displaystyle\int_{o=0}^{(\pi/2)} S(\vec{U})\, do}.$$

We notice that the value of $S(\vec{U})$ does not depend on θ as a result of the spherical symmetry. The exact expression of the a priori probability density of apparition of a direction $\vec{U}$ is given by the normalization of the function S, that is, D, because the allowed zone is the total Gaussian sphere, $S(\vec{U}) = 1$. If $\vec{U}$ points out the image, $S(\vec{U}) = (\Delta\theta')/\pi$. The width of the allowed zone depends also on the position in θ of $\vec{U}$: there is no more symmetry with respect to the z axis. Calculus is developed in [15].

Appendix II
A Priori Probability Density of Vertical $\vec{V}$

We have made the calculation only in the case of a circular retina. The $\vec{U}$ contributing to the apparition of a vertical $\vec{V}$ are located on the great circle perpendicular to $\vec{V}$. Thus, the repartition of $\vec{V}$ is related to the repartition of $\vec{U}$ by

$$R(\vec{V}) = \frac{2}{\pi} \int_{C/4} P(\vec{U})\, ds, \qquad \vec{V} \begin{bmatrix} o_{V} \\ \theta_{V} \end{bmatrix}.$$

$C/4$ is a quarter of the great circle that is perpendicular to direction $\vec{V}$. Thus the equation shown at the bottom of the next page applies. The integral involved in the two formulas is an elliptic integral and cannot be calculated explicitly. After normalization of the function $R(\vec{V})$, we obtain the exact expression of the a priori probability density of a vertical. It is now clear that the vertical lines near the (x, y) plane are more probable (see Fig. A2). We have also verified these computations by simulation (in the same condition as above; see Fig. A2).

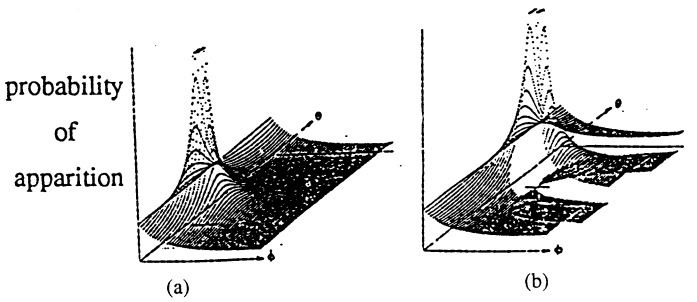

Fig. A3. Impulse response for a circular retina of diameter 512 pixels, pixel size of $25 \times 25\,\mu$m, and a focal length of 50 mm, for $\vec{U}_0$. (a) Inside the projection of the retina. (b) Outside the projection of the retina.

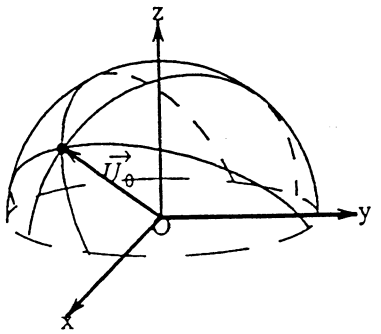

Fig. A4. Repartition of the great circles for an infinite retina.

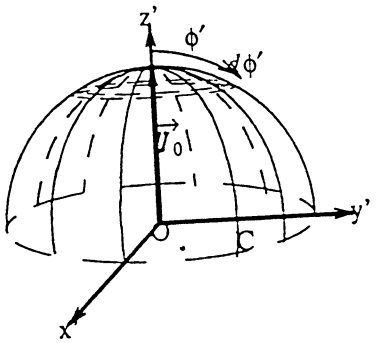

Fig. A5. Centering of the problem and calculation of $Q(o')$.

Appendix III
Impulse Response

The aim of this section is to find the shape of the accumulator on the Gaussian sphere, when a "signal" is in the image, that is, when a vanishing direction is represented in the image of the 3D scene. Let $\vec{U}_0$ denote a direction of the scene infinitely represented by segments, that is, through each point of the image, passes a line, projection of a 3D line parallel to $\vec{U}_0$.

A. Hypothesis of an Infinite Retina

Suppose that the image plane is not restricted by a retina, and the accumulator is thus composed by an infinity of great circles of the Gaussian sphere consisting each other in $\vec{U}_0 \begin{bmatrix} o_0 \\ \theta_0 \end{bmatrix}$. This configuration on the entire Gaussian sphere is symmetrical with respect to $\vec{U}_0$. Thus,

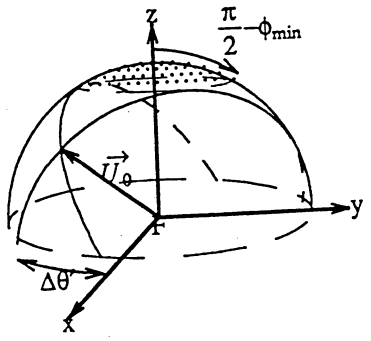

Fig. A6. Repartition of the great circles for a finite retina.

we can calculate it easily in a new coordinate frame (x', y', z'), where z' is parallel to $\vec{U}_0$.

The density Q depends only on o' (angle of the vector with the z' axis) in this new frame. According to the calculation proposed by Brown [4], we can assume that $Q(o')$ is proportional to the length of dA of a circle arc C cut by a ring defined by o' and do', divided by the surface of this ring: $Q(o') = dA/(2\pi \sin o' do')$. Thus $Q(o') = 1/(\pi \sin o')$.

In the original coordinate frame, o' is the angle between $\vec{U}_0$ and $\vec{U}$, that is, $\cos o = \vec{U}_0 \cdot \vec{U}$. This leads to

$$Q(\vec{U}) = \frac{1}{\pi\sqrt{1 - (\vec{U}_0 \cdot \vec{U})^2}}.$$

B. Case of a Finite Retina

We will take the hypothesis of a circular retina for simplicity. The segments (or lines) representing $\vec{U}_0$ exist if and only if their $\vec{N}$ vector (normal to their interpretation plane) is such that $o_N \in [o_{\min}, \pi/2]$. Therefore, the great circles covering the Gaussian sphere are not distributed uniformly, and the impulse response is no more strictly symmetrical around $\vec{U}_0$. The fact that $\vec{N}$ is limited is related directly to the fact that the great circles must pass through the projection of the retina on the Gaussian sphere. This projection is a portion of the sphere centered on the z axis and whose aperture angle is $(\pi/2) - o_{\min}$ (Fig. A6).

If $o_0 \leq (\pi/2) - o_{\min}$, $\vec{U}_0$ is in the projection of the retina. Because all the circles cross each other in $\vec{U}_0$, they all cut the projection of the retina. The same calculation holds that was made in the case of an infinite retina.

If $o_0 \geq (\pi/2) - o_{\min}$, only the great circles cutting the projection of the retina are crossing each other in $\vec{U}_0$. Let $\Delta o'$ be the aperture angle of the allowed zone (i.e., the angle between the two delimiting planes), which is symmetrical with respect to the $(\vec{z}, \vec{U}_0)$ plane. $\Delta o' = \arcsin[\cos o_{\min} / \sin o_0]$.

For $\vec{U}$ not belonging to the so-defined "part of sphere," we have $Q(\vec{U}) = 0$.

If $\vec{U}$ is such that $\theta' \in [0, \Delta\theta'] \cup [\pi - \Delta\theta', \pi + \Delta\theta'] \cup [2\pi - \Delta\theta', 2\pi]$, θ' is measured in the plane perpendicular to $\vec{U}_0$ and from the projection of $\vec{z}$ on this plane. Thus $Q(\theta', o')$ is independent of θ':

$$Q(\theta', o') = \frac{dA}{4\Delta\theta' \sin o' do'} = \frac{1}{2\Delta\theta' \sin o'}.$$

$$R(\vec{V}) = \begin{cases} \dfrac{4}{\pi^2}\displaystyle\int_0^{\pi/2} \arcsin\left[\frac{\cos o_0}{\sqrt{1 - \sin^2 s \sin o_V}}\right] ds, & \text{if } o_V \leq o_0 \\ \dfrac{4}{\pi^2}\displaystyle\int_0^{\arcsin(\sin o_0 / \sin o_V)} \arcsin\left[\frac{\cos o_0}{\sqrt{1 - \sin^2 s \sin o_V}}\right] ds + \frac{2}{\pi}\left(\frac{\pi}{2} - \arcsin\left(\frac{\sin o_0}{\sin o_V}\right)\right), & \text{if } o_V \geq o_0. \end{cases}$$

The shape of this impulse response is presented in Fig. A3. Its expression in the original frame is

- If $o_0 \leq \frac{\pi}{2} - o_{\min}$ $\quad Q(\vec{\ell}) = \dfrac{1}{\pi\sqrt{1-(\vec{\ell}_0 \cdot \vec{\ell})^2}}$ (A1)

- If $o_0 \geq \frac{\pi}{2} - o_{\min}$ (A2)

$$Q(\vec{\ell}) = \begin{cases} 0. & \text{if } |\sin o \sin(\theta - \theta_0)| \geq \dfrac{\cos o_{\min}}{\sin o_0}\sqrt{1-(\vec{\ell}_0 \cdot \vec{\ell})^2} \\ \dfrac{1}{2 \arcsin\left[\dfrac{\cos o_{\min}}{\sin o_0}\right]\sqrt{1-(\vec{\ell}_0 \cdot \vec{\ell})^2}}. & \text{if } |\sin o \sin(\theta - \theta_0)| \leq \dfrac{\cos o_{\min}}{\sin o_0}\sqrt{1-(\vec{\ell}_0 \cdot \vec{\ell})^2} \end{cases} \quad \text{(A3)}$$

Remark: By integrating (A1), (A2), and (A3) for all the $\vec{\ell}_0$ directions, we obtain exactly (1), which expresses the bias:

$$\int_{\theta=0}^{\pi/2}\int_{o=0}^{2\pi} Q(\vec{\ell}) = S(\vec{\ell}).$$

This result is logical: The bias and the impulse response are two exhibitions of one phenomenon.

REFERENCES

[1] N. Badler, "Three-dimensional motion from two-dimensional picture sequences," in *Proc. 2nd Int. Joint Conf. Pattern Recognition*, Copenhagen, Denmark, Aug. 13–15, 1974, pp. 157–161.

[2] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[3] S. T. Barnard, "Interpreting perspective images," *Artificial Intell.*, vol. 21, 1983, pp. 435–462.

[4] C. M. Brown, "Inherent bias and noise in the Hough transform," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, no. 5, pp. 493–501, Sept. 1983.

[5] P. V. C. Hough, "Method and means of recognizing complex patterns," U.S. Patent 3,069,654,18, Dec. 1962.

[6] A. Iannino, "Hough transform theory and image processing experiments," Ph.D. dissertation, Stevens Institute of Technology, 1979.

[7] J. Illingworth and J. Kittler, "A survey of the Hough transform," *Comput. Vision, Graphics, Image Processing*, vol. 44, pp. 87–116, 1988.

[8] B. Kamgar-Parsi and B. Kamgar-Parsi, "Evaluation of quantization error in computer vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, pp. 929–940, Sept. 1989.

[9] J. Kender, "Shape from texture: An aggregation transform that maps a class of textures into surface orientation," in *Proc. Int. Joint Conf. on Artificial Intelligence*, Tokyo, Japan, Aug. 20–23, 1979, pp. 475–480.

[10] H. Li, M. A. Lavin, and R. J. Le Master, "Fast Hough transform: A hierarchical approach," *CVGIP 36*, 1986, pp. 139–161.

[11] H. Li and M. A. Lavin, "Fast Hough transform based on bintree data structure," *CVPR 86*, pp. 640–642.

[12] J. Lopez-Krahe and R. Villata, "Transformation de Hough discrète et bornée. Applications à la détection de primitives rectangulaires," *MARI 87*, Paris, La villette, Mai 18–22, 1987, pp. 42–48.

[13] J. Lopez-Krahe and P. Pousset, "Transformée de Hough discrète et bornée. Application à la détection de droites parallèles et du réseau routier," *Colloque TIPI 88, Traitement du Signal*, vol. 5, no. 4, pp. 281–289, 1988.

[14] E. Lutton, H. Maitre, and J. Lopez-Krahe, "Determining vanishing points with Hough transform," in *Proc. SPIE Visual Communications and Image Processing '90*, Lausanne, Switzerland, Oct. 2–4, 1990, pp. 420–430.

[15] E. Lutton, "Reconnaissance du point de prise de vue d'une photographie à partir d'un modèle de scène," Thèse de Télécom Paris, 90E005, May 1990.

[16] M. J. Magee and J. K. Aggarwal, "Determining vanishing points from perspective images," *CVGIP 26*, 1984, pp. 256–267.

[17] H. Maitre, "Un panorama de la transformation de Hough," *Traitement du Signal*, vol. 2, no. 4, pp. 305–317, 1985.

[18] H. Maitre, "Contribution to the prediction of performances of the Hough transform," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 5, pp. 669–674, Sept. 1986.

[19] L. Quan, R. Mohr, and E. Thirion, "Generating the initial hypothesis using perspective invariants for a 2D image and 3D model matching," in *Proc. 9th Int. Conf. Pattern Recognition*, Rome, Italy, 1988, pp. 872–874.

[20] L. Quan and R. Mohr, "Determining perspective structures using hierarchical Hough transform," *Pattern Recogn. Lett.*, vol. 9, 1989, pp. 279–286.

[21] H. A. Sedgwick, "LAYOUT2: A production system modeling visual perspective information," in *Proc. 1st Int. Conf. Computer Vision*, London, June 8–11, 1987, pp. 223–229.

[22] S. D. Shapiro, "Transformations for the computer detection of curves in noisy pictures," *CVGIP 4*, 1975, pp. 328–338.

[23] S. D. Shapiro and A. Iannino, "Geometric constructions for predicting Hough transform performance," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, no. 3, pp. 310–317, July 1979.

[24] I. D. Svalbe, "Natural representations for straight lines and the Hough transform on discrete arrays," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, pp. 941–950, Sept. 1989.

Morphological Filtering as Template Matching

Ronald Jones and Imants Svalbe

Abstract—**Binary morphological operations using single and multiple structuring elements are implemented using look-up table (LUT) driven templates. Many complex operations can be implemented in one pipeline processing cycle for 3*3 regions of support and in four or five cycles for 5*5 regions of support. The basis representation of the operations is used to specify the required templates.**

Index Terms—**Binary morphology, morphological basis decomposition, parallel processing, template matching.**

I. INTRODUCTION

Morphologic filtering traditionally combines serial erosions and dilations to process images in a shape-sensitive way [1], [2]. It has long been established that any translation-invariant increasing operation (such as an open or close) has an equivalent basis representation. An operation Ψ is then expressed through a single-step process as a union of erosions by the basis set members [3], [4]:

$$\Psi(X) \equiv \bigcup_{A} X \div A. \quad (1)$$

The basis set $\{A\}$ lists all of the local data structures for any input image X that will cause pixels to be on or off in the output image. Simple algorithms exist [5]–[8] to derive the basis representation for many operations.

Although the data level description of the basis representation offers insight into the effect of an operation, actually implementing operations as a union of many sequential individual erosions is cumbersome. For a large structuring element, the basis set will usually have many members because the set must exemplify all possible changes the operation can produce. The size or support of the basis members also increases with the size of the structuring element and

Manuscript received December 30, 1991; revised May 18, 1993. Recommended for acceptance by Associate Editor E. Delp.

The authors are with the Department of Physics, Monash University, Clayton 3168, Australia.

IEEE Log Number 9214455.