INFORMATION AND CONTROL 30, 169–186 (1976)

On the Number of Sides Necessary for Polygonal Approximation of Black-and-White Figures in a Plane*

JAMES R. ELLIS, JR. AND MURRAY EDEN

Department of Electrical Engineering and Computer Science
and
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A bound on the number of extreme points or sides necessary to approximate a convex planar figure by an enclosing polygon is described. This number is found to be proportional to the fourth root of the figure's area divided by the square of a maximum Euclidean distance approximation parameter.

An extension of this bound, preserving its fourth root quality, is made to general planar figures. This is done by decomposing the general figure into nearly convex sets defined by inflection points, cusps, and multiple windings.

A procedure for performing actual encoding of this type is described. Comparisons of parsimony are made with contemporary figure encoding schemes.

INTRODUCTION

The encoding of geometric figures in the Euclidean plane was a matter of little interest until fast, large-scale digital computation became commonly available. Soon after cathode-ray tube scanning for pictorial sampling and display became practical, it was noted that representations of black-and-white figures that had subjectively simple characteristics did not require specification of the intensity and location of each point. Freeman (1961) was the first to publish work considering the information lossless encoding of sampled images by contour representations. Most subsequent publications have been concerned with the determination of properties of sampled figures. Minsky and Papert (1972) and Duda and Hart (1973) have published bibliographies relevant to this area.

There have been a number of primarily experimental studies on photographs of natural objects. These include theses by Walpert (1970), Young (1966), Graham (1967), and Pan (1962). These results and others have shown

* This work was supported in part by the Joint Services Electronics Program Contract No. DAA BO7-74-C-0630.

that for most real pictures, there are a variety of heuristic techniques that can achieve significant data compression either without loss of information or without loss of subjective quality. Progress in this area is reviewed by Huang, Schreiber, and Tretiak (1971).

Recently, works have begun to consider some of the theoretical aspects of digital picture coding. Freeman and Glass (1969) have considered a contour "energy-related" procedure for determining adequacy of approximation. Montanari (1968, 1969) has considered finding skeletons as a binary picture compression device. Sklansky and his collaborators (1970, 1972a,b,c) have developed the minimum-perimeter polygon as an unambiguous descriptor of a cellular or sampled image. Montanari (1970) has studied the limiting properties of the Freeman and Sklansky representations in considerable detail. He found it necessary to define at least two types of equivalence for digitization schemes in order to find a unique and meaningful solution to the minimization problems that both the Freeman and Sklansky procedures require.

The questions of bounds on the information required for these representations has been discussed by Eden (1960, 1968) and Bolour and Cover (1972). The former shows that for nearly convex "blobs" the number of bits necessary to encode a figure of k cells is approximately $k + 4k^{1/2}$. The latter states the strong result that a convex figure of perimeter L can be ϵ-approximated by about $(\pi(L/\epsilon))^{1/2}$ straight line segments. As we shall see, ϵ is not the same as, but is related to, the sampling interval.

This paper considers bounds on the number of points, and thus, on the number of bits, necessary to encode arbitrary geometric figures. The bounds are related to the figure's area for convex figures and to derived areas for figures with concavities.

An Area-Related Bound on the Polygonal Approximation of Convex Sets

This section considers one aspect of the coding of convex sets in two dimensions. It demonstrates that any convex set with area A can be ϵ-approximated by a convex polygonal hull with no more than $o(A/\epsilon^2)^{1/4}$ sides or extreme points. It is shown that the number n of sides in a polynogal ϵ-approximation is bounded from above by

$$n \leqslant m_A = K(m_A)\,\pi^{3/4}(A/\epsilon^2)^{1/4} + 3,$$

where $K(m_A)$ is a monotonically decreasing function of m_A, ranging in value from $3^{9/8}/\pi^{3/4}$ ($\simeq$1.46) for $m_A = 6$, to an asymptotic value of 1.

This result may be compared with a previous result by Bolour and Cover (1972) that the bound on the number of sides is related to the figure's perimeter L by

$$n \leqslant m_L = \pi(L/2\pi\epsilon + 1)^{1/2}.$$

For a circle of radius R, we have

$$\pi^{3/4}(A/\epsilon^2)^{1/4} = \pi(R/\epsilon)^{1/2} = \pi(L/2\pi\epsilon)^{1/2}.$$

Since it is well known that a circle is the planar figure having the largest ratio of area to perimeter, it follows that for (A/ϵ^2) and $(L/2\pi\epsilon)$ large, $m_A < m_L$. For an ellipse with semiaxes a and c, $a < c$, the asymptotic behavior is

$$m_A/m_L \sim (\pi/2)^{1/2}(a/c)^{1/4}.$$

The Bolour and Cover result, in turn, may be compared with the Freeman chain approach (Freeman, 1961) of connecting $m_E = L/\epsilon$ adjacent samples at intervals of ϵ along the perimeter of the set. The latter is a great improvement itself over the idea of specifying the values of an entire grid large enough to cover the figure with precision ϵ!

The approach taken here is based conceptually on that used by Bolour and Cover. Specifically, a circumscribing polygon that satisfies the conditions imposed is constructed. Thus, the existence of a solution is explicitly shown. Then an inscribed polygon, which touches the circumscribing polygon and the coded periphery at the tangency points of the latter two, is used in a series of dominance relations to establish a bound on a trigonometric function of the number of sides. The positive second derivative nature of the functions found is used to establish the result for normal or "slowly turning" regions of the periphery and to set the limits of this behavior. It is shown that no more than three regions of a curve can be "sharp corners" or abnormal in this sense.

Before turning to the technical aspects of ϵ-approximation, we define the term. We say that a set S ϵ-approximates a set $\mathbf{S}$ if $\mathbf{S} \subseteq S$, and $d(S, \mathbf{S}) \leqslant \epsilon$; where $d(S, \mathbf{S}) \equiv \max_s \min_{\mathbf{s}} \| s - \mathbf{s} \|$, s and $\mathbf{s}$ are members of S and $\mathbf{S}$, respectively, and $\| \; \|$ is the Euclidean metric in the plane. Figure 1 illustrates this concept.

For the purpose of proof, we will inscribe a polygon in the convex set. Then, a central point may be picked for the purpose of forming disjoint sets covering the polygon. One such wedge is illustrated in Fig. 2. It should be obvious that any point interior to the convex set is sufficient for the arguments that follow, although the centroid of the set or some point near it is intuitively satisfying as a center for the wedges.

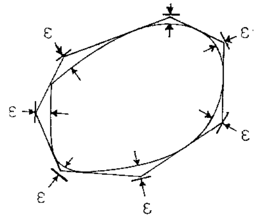

FIG. 1. Illustration of an ε-approximation of a convex set by a polygon.

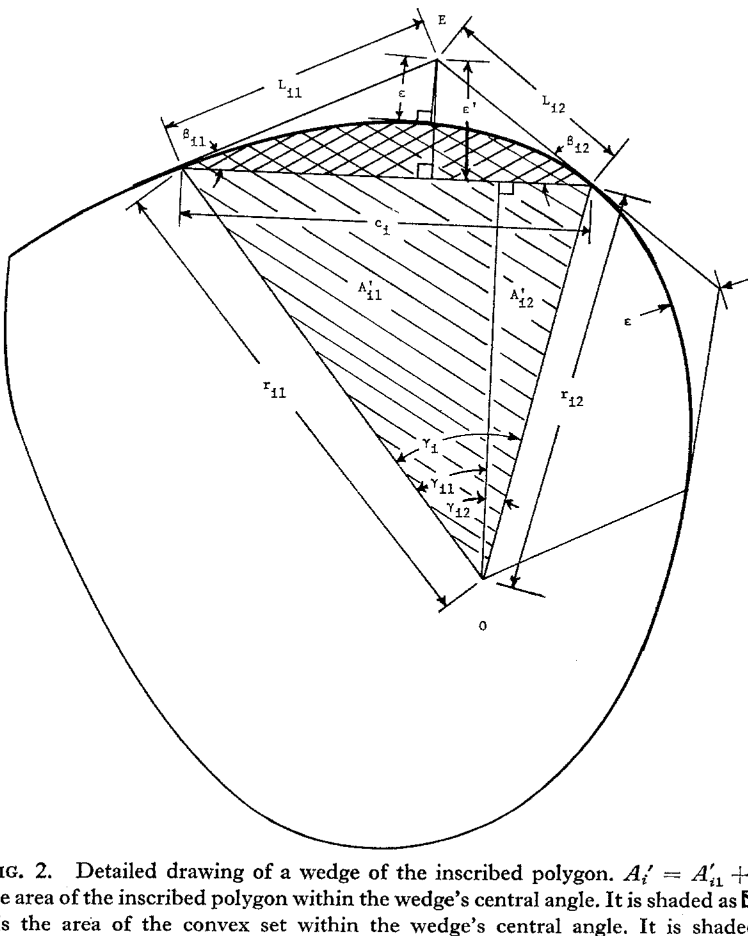

FIG. 2. Detailed drawing of a wedge of the inscribed polygon. $A_i' = A_{i1}' + A_{i2}'$ is the area of the inscribed polygon within the wedge's central angle. It is shaded as ▧. A_i is the area of the convex set within the wedge's central angle. It is shaded as ▧ ∪ ▩. 0 is the central point, T_{i1} and T_{i2} are tangency points of the set and its circumscribed polygon, and E is the extreme point of the intersection of these tangent lines. Other line segments and angles should be easily interpretable.

The normal distance from the extreme point to the chord of the inscribed polygon may be used to relate the outer angles and tangential distances:

$$\epsilon' = L_{i1} \sin(\beta_{i1}) = L_{i2} \sin(\beta_{i2}). \tag{1}$$

The inscribed chord's length is

$$c_i = L_{i1} \cos(\beta_{i1}) + L_{i2} \cos(\beta_{i2}). \tag{2}$$

Using the normal distance relation to rewrite the chord length in terms of one distance parameter, we obtain:

$$c_i = L_{i1}(\cos(\beta_{i1}) + (\sin(\beta_{i1})/\sin(\beta_{i2})) \cos(\beta_{i2}) \tag{3a}$$

$$= L_{i1} \sin(\beta_{i1})(\cot(\beta_{i1}) + \cot(\beta_{i2})). \tag{3b}$$

Similar relationships hold for the radial distances r_i and the central angles γ_i. Because of the definitions used, sines and cosines are interchanged in these similar equations:

$$c_i = r_{i1} \cos(\gamma_{i1})(\tan(\gamma_{i1}) + \tan(\gamma_{i2})). \tag{3c}$$

We can now write the area of the inscribed polygon in terms of a single distance and trigonometric functions of these angles.

$$A_i' = A_{i1}' + A_{i2}', \tag{4a}$$

$$= (1/2)\, r_{i1} \cos(\gamma_{i1})(L_{i1} \cos(\beta_{i1}) + L_{i2} \cos(\beta_{i2})), \tag{4b}$$

$$= \frac{(L_{i1} \sin(\beta_{i1})(\cot(\beta_{i1}) + \cot(\beta_{i2}))^2)}{(2(\tan(\gamma_{i1}) + \tan(\gamma_{i2})))}. \tag{4c}$$

Since

$$\epsilon' \geqslant \epsilon, \tag{5r1}$$

using (1), we obtain an inequality relating A_i' and ϵ:

$$A_i' \geqslant \epsilon^2(\cot(\beta_{i1}) + \cot(\beta_{i2}))^2/(2(\tan(\gamma_{i1}) + \tan(\gamma_{i2}))). \tag{5}$$

The right-hand side of this inequality has positive second derivatives with respect to all angles for which

$$0 \leqslant \beta_{i1}, \beta_{i2}, \gamma_{i1}, \gamma_{i2} \leqslant \pi/2, \tag{5r2}$$

and

$$|\gamma_{i1} + \gamma_{i2}| \leqslant \pi/2. \tag{5r3}$$

We will use the fact that if $f(x)$ has nonnegative second derivative,

$$\sum \lambda_i f(x_i) \geqslant f\left(\sum \lambda_i x_i\right), \tag{5up1}$$

where $0 \leqslant \lambda_i \leqslant 1$; $\sum \lambda_i = 1$. For reference, we label this Lemma 1. One clear example is

$$(f(a) + f(b))/2 \geqslant f(a + b)/2. \tag{5up2}$$

Inequality (5up1) follows by obvious induction.

Let us consider separately the wedges associated with extreme points at which the positive second derivative restriction (partial derivatives, to be correct) is true—that is, those that satisfy (5r3)—and those at which it is not.

For the former, or normal segments, use of Lemma 1 gives

$$\begin{aligned} &\sum (\lambda_i(\cot(\beta_{i1}) + \cot(\beta_{i2}))^2/(\tan(\gamma_{i1}) + \tan(\gamma_{i2}))) \\ &\qquad \geqslant \left(\cot\left(\sum \lambda_i(\beta_{i1} + \beta_{i2})/2\right)\right)^2 \Big/ \left(\tan\left(\sum \lambda_i(\gamma_{i1} + \gamma_{i2})\right)\right), \end{aligned} \tag{6}$$

when (5r3) is satisfied (i.e., $0 \leqslant \gamma_{i1} + \gamma_{i2} \leqslant \pi/2$, and where

$$0 \leqslant \lambda_i \leqslant 1; \tag{6a}$$

$$\sum \lambda_i = 1. \tag{6b}$$

Summations are over all i, assumed to be $1 \leqslant i \leqslant m_1$. Specifically, we will use

$$\lambda_i = 1/m_1 . \tag{6c}$$

Since

$$\sum (\beta_{i1} + \beta_{i2}) \equiv B \leqslant 2\pi, \tag{7a}$$

and

$$\sum (\gamma_{i1} + \gamma_{i2}) \equiv G \leqslant 2\pi, \tag{7b}$$

we can obtain an algebraic bound on the number of sides of the polygons under the positive second-derivative restriction. Using relations (5), (6), and (7), we get

$$A \geqslant \sum A_i' \tag{8a}$$

$$\geqslant \sum (\epsilon^2/2)((\cot(\beta_{i1}) + \cot(\beta_{i2}))^2/(\tan(\gamma_{i1}) + \tan(\gamma_{i2}))). \tag{8b}$$

$$2A/\epsilon^2 \geqslant \sum (2^2(\cot((\beta_{i1} + \beta_{i2})/2))^2/(\tan(\gamma_{i1}) + \tan(\gamma_{i2}))), \tag{8c}$$

$$\geqslant \sum (2^2(\cot((\beta_{i1} + \beta_{i2})/2))^2/\tan(\gamma_{i1} + \gamma_{i2})). \tag{8d}$$

$$A/\epsilon^2 \geqslant 2m_1 \left(\left(\cot\left(\sum (\beta_{i1} + \beta_{i2})/2m_1\right)\right)^2 \Big/ \tan\left(\sum (\gamma_{i1} + \gamma_{i2})/m_1\right)\right), \tag{8e}$$

$$= 2m_1((\cot(B/2m_1))^2/\tan(G/m_1)), \tag{8f}$$

$$\geqslant 2m_1(\cot(\pi/m_1))^2 \cot(2\pi/m_1), \tag{8g}$$

$$= 2m_1(m_1/\pi)^2(m_1/2\pi)(\cot(\pi/m_1)(\pi/m_1))^2(\cot(2\pi/m_1)(2\pi/m_1)). \tag{8h}$$

Collecting terms, we obtain

$$m_1^4 \leqslant \pi^3(A/\epsilon^2)(\tan(\pi/m_1)/(\pi/m_1))^2(\tan(2\pi/m_1)/(2\pi/m_1)). \quad (9a)$$

Then, using $(\tan(x)/x)$ bounds, we get

$$m_1 \leqslant ((3^{3/2}/\pi)^3\pi^3(A/\epsilon^2))^{1/4} = 3^{9/8}(A/\epsilon^2)^{1/4}, \qquad \text{for} \quad m_1 \geqslant 6; \quad (9b)$$

$$\lim_{m_1 \to \infty}(m_1) = \pi^{3/4}(A/\epsilon^2)^{1/4}. \quad (9c)$$

If a wedge falls into the second class, we can always add another side to the circumscribing polygon wherever a central angle of more than $\pi/2$ is encountered. This is illustrated in Fig. 3. Since a convex polygon has a total

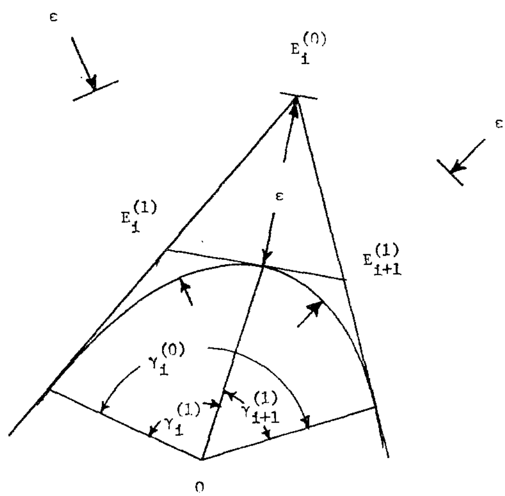

FIG. 3. Illustration of the addition of a side to the circumscribing polygon which ϵ-approximates a convex region, but which gives a central angle of greater than $\pi/2$. Clearly, $E_i^{(1)}$ and $E_{i+1}^{(1)}$ are less than ϵ from the region's perimeter, and $\gamma_i^{(1)}$ and $\gamma_{i+1}^{(1)}$ can be made less than $\pi/2$, by repeated applications if necessary.

central angle of 2π, there can be at most three angles greater than $\pi/2$. Thus, we have shown that the bound on the number of members of the second class is

$$m_2 = 3. \quad (10)$$

The upper bound we are seeking is just the total number of elements in these two classes. This is clearly

$$m = m_1 + m_2 . \quad (11)$$

Substituting (9a) for m_1, we have Theorem 1:

$$m \leqslant \pi^{3/4} \left[(A/\epsilon^2) \frac{(\tan(\pi/m_1))^2}{(\pi/m_1)^2} \cdot \frac{\tan(2\pi/m_1)}{(2\pi/m_1)} \right]^{1/4} + 3. \tag{12a}$$

As m becomes very large,

$$m \simeq \pi^{3/4}(A/\epsilon^2)^{1/4}. \tag{12b}$$

Two items seem worth noting here. First, if we are interested in ϵ-approximating a convex figure that is roughly 1000×1000 ϵ-units in area, this bound says that no more than about 90 sides are necessary for construction of the polygon.

Second, intuition suggests that the limiting case as $\epsilon \to 0$ for polygonal approximation is the circle. Which convex figure of area A actually requires a maximal number of ϵ-approximating edges is an open question. For a circle of radius R, both this bound and the perimeter bound give

$$m \simeq \pi(R/\epsilon)^{1/2}. \tag{13a}$$

Straightforward construction gives

$$n_{\text{circle}} \simeq \pi(2^{-1/2})(R/\epsilon)^{1/2}. \tag{13b}$$

The difference was finally traced to inequality (5r1), where in the case of a circle

$$\epsilon' \simeq 2\epsilon. \tag{13c}$$

Bolour and Cover's proof has a similar generous overbound.

The bound is a very smooth, monotonically increasing function of A/ϵ^2. Because it must dominate many discrete cases, where n must make integral jumps at certain values of ϵ for a given A, it is not clear that the lower bound indeed has an overgenerous factor of $\sqrt{2}$. In particular, the ϵ-approximation of regular or nearly regular n_1-sided polygons by n_2-sided polygons, where $n_1 < n_2$ and n_1 is relatively prime to n_2, sometimes requires a larger n_2 than a circle of the same area. For example, the minimum ϵ for which it is possible to find a triangular approximation is smaller for a circle than for a square.

Extension of the Bound to Nonconvex Sets

To extend the result obtained previously for convex sets, we may note first of all that for simply connected sets that are fairly simple or "blob-like," we can decompose the boundary by finding its inflection points. Then we can apply the bound to the convex and nearly convex sets (to be defined below) formed by the portion of the boundary between two adjacent inflection points and the straight line segment between these points.

If there are N_{IP} ($N_{\mathrm{IP}} \geqslant 2$) inflection points and cusps on a figure's boundary, and the figure has a convex hull with area $\mathbf{A}$, then clearly there are no more than N_{IP} sets, and each has an area no greater than $\mathbf{A}$. So the number of sides necessary for such a figure is bounded by

$$n \leqslant m_B \leqslant N_{\mathrm{IP}} m_{\mathrm{NC}}(\mathbf{A}), \tag{14}$$

where m_{NC} is the bound for nearly convex regions, defined in Eq. (16). An example of this decomposition is given in Fig. 4.

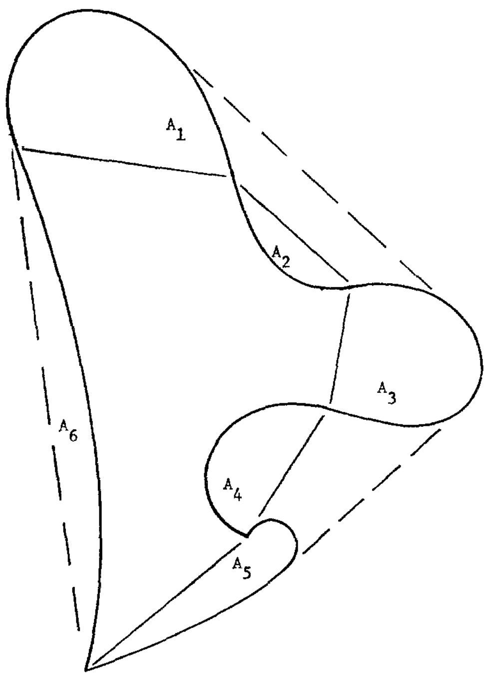

FIG. 4. Illustration of the decomposition of a nonconvex set into convex (and nearly convex) subsets. Convex hull boundaries are dashed lines where different from original set boundary.

A better bound can be generated with relatively little work using the same ideas. The boundary is traced, and the convex break points are found. Then

$$m_B \leqslant \sum (K_A(m_{1i})\, \pi^{3/4}(A_i/\epsilon^2)^{1/4} + 3), \tag{15}$$

where, as before, m_{1i} is the number of sides in the smooth section bound, and $K_A(j) = (\tan(j)/j)^2(\tan(2j)/2j)$. Each of the areas in this sum is associated with one of the smaller sets indicated in Fig. 4. If any of these sets is not convex, then the nearly convex bound should be used as necessary in Eq. (15).

One definition of such blob-like regions is that at least one point exist for which the winding number of a point traversing the boundary goes from 0 monotonically to 1. When this winding number condition is not satisfied, one encounters some difficulty defining a minimal number of convex regions that include all of the boundary. This occurs, for example, in Fig. 5.

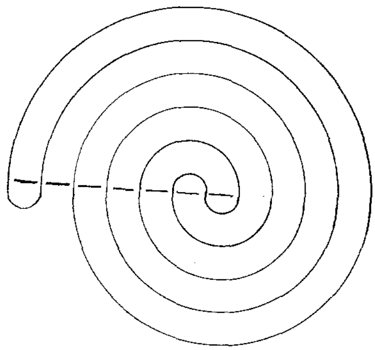

FIG. 5. A spiral region illustrating a large winding number situation.

Note that a figure of this sort can be drawn with only two inflection points. In this case, a line segment drawn between the inflection points results in a number of nearly convex sets similar to the one shown in Fig. 6a.

For such a region, the arguments made in the convex-set coding bound proof hold with slight modifications. First, in Eq. (7a), B could be somewhat greater than 2π, say 3π, but no greater. Second, there might be two more sides for the extra turning angle segment. In this case, a factor of $(3\pi/2\pi)$ is introduced into the $\cot^2(B/2m_1)$ term in Eq. (8f), so that the coefficient in m_1 goes from $\pi^{3/4}$ to $\pi^{3/4}(3/2)^{1/2}$, and m_2 goes from 3 to 5 in Eq. (10).

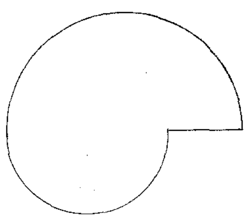

FIG. 6a. An example of a nearly convex component of a spiral.

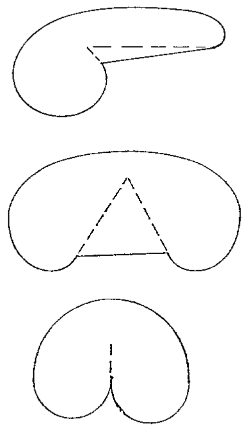

FIG. 6b. Illustration of a few nearly convex regions. For the centers indicated by dotted line intersections, both the central and peripheral turning angles are monotonically increasing (or decreasing) and satisfy the integral limits $B \leqslant 3\pi$ (peripheral turning angle), and $G \leqslant 2\pi$ (central turning angle).

Changing the limits allows one to avoid consideration of the value of the angles of the boundary crossing the straight line break segments. Regions such as the one illustrated in Fig. 6b are included. We will call regions satisfying these less restrictive conditions "nearly convex." For purposes of reference, we will write the ϵ-coding bound for such regions as Theorem 2:

$$m_{\mathrm{NC}} \leqslant K_A(m_1)\, \pi^{3/4}(3/2)^{1/2}(A/\epsilon^2)^{1/4} + 5. \tag{16}$$

It seems that such a change in the limits allows one to use only inflection points and cusps as critical points, and thus, to find the coding limits with purely local boundary operations.

If the smallest integer greater than or equal to the maximum winding number is N_w, then there may be $2N_w$ regions of this type. There also may be two regions at the ends of the "loop" that must be bounded separately. So, with **A** the convex hull area, as previously, we can write a bound for spiral regions as Corollary 2A:

$$m_w \leqslant 2(N_w + 1)(3^{9/8}(3/2)^{1/2}(\mathbf{A}/\epsilon^2)^{1/4} + 5). \tag{17}$$

If, as before, we identify the individual areas, a stronger bound may be obtained at the cost of more computation. In this case, we have Corollary 2B:

$$m_w \leqslant \sum (K_A(m_{1i})(3/2)^{1/2}\pi^{3/4}(A_i/\epsilon^2)^{1/4} + 5), \tag{18}$$

where the symbols have the same meaning as before. If the sets into which the original figure is decomposed are selected by the use of adjacent critical points, where critical points are defined as inflection points, cusps, or points where the tangent vector of the boundary has swept through 2π without encountering an inflection point or cusp, then Corollary 2B can be used for computing the bound for any closed figure in that form.

A slightly different approach is the following. If we pick a central point more restrictively, so that the winding number of our boundary is monotone (as it is in the case of a point near the center of a spiral), we can then define "area" as the integral swept over by a line segment from our central point to the periphery. Thus, some areas of the figure (and its convex-hull complement) are counted multiply, as in Fig. 7.

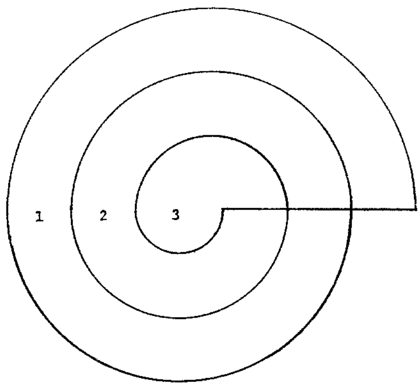

FIG. 7. Illustration of multiply counted areas in a spiral decomposition.

In this situation, we can set the limits on G as $G \leqslant 2\pi N_w$, and, assuming the worst case on the periphery, on B as $B \leqslant 2\pi N_w + \pi$. B and G are defined in Eqs. (7a) and (7b). Since the total turning angle of the boundary is also bounded by B, m_2, the sharp corner bound, becomes

$$m_2 \leqslant B/(\pi/2) - 1 = 4N_w + 1. \tag{19}$$

Thus, the ϵ-approximation polygonal bound for a figure having monotone winding number N_w and area A' is Corollary 2C:

$$m_w \leqslant \pi^{3/4}(N_w + 1/2)^{1/2} N_w^{1/4} (A'/\epsilon^2)^{1/4} K_A(m_1) + 4N_w + 1, \tag{20a}$$

$$\leqslant 3^{9/8} N_w^{1/4} (N_w + 1/2)^{1/2} (A'/\epsilon^2)^{1/4} + 4N_w + 1. \tag{20b}$$

This bound is better than the previous one only in the case where the area in the spiral is concentrated near its center, except for the factor of $(\frac{3}{2})^{1/2}$. That factor is about $\frac{1}{4}$ better, which can be nontrivial for large m_w.

Appropriate application of Theorem 2, Corollary 2A, and Corollary 2B should give polygonal approximation bounds for any figure of interest in pattern recognition or image coding.

A Procedure for Approaching These Bounds

We base our procedure on the Sklansky–Chazin–Hansen (1972) (SCH) algorithm. First, we find the minimum-perimeter polygon (MPP) of a cellular complex in the SCH sense. Call this ∂J. Then, shift the boundary of the MPP normally outward by ϵ and connect adjacents ends of its expansion by straight line segments. This gives us a cellular boundary of the type shown in Fig. 8, which we call $\partial(J + \epsilon)$.

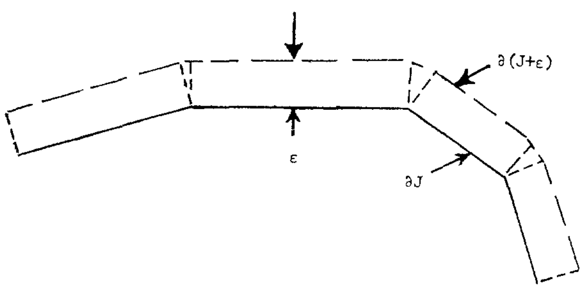

FIG. 8. Illustration of a boundary expansion for an MPP.

We will assume essentially infinite-precision arithmetic, and sampling and extreme-point specification on a $b\epsilon$-net. The parameter b will need to be less than 1 for the following arguments to apply, most conveniently $b = 2^{-n}$, where n is a positive integer. We will first assume a convex cellular image.

We will fit line segments into the tube formed by ∂J and $\partial(J + \epsilon)$. The extreme points will be $b\epsilon$ lattice points close to (or on) the outer boundary $\partial(J + \epsilon)$. The vertices of ∂J will be constraint points through which or close to which the line segments will pass. Note that we are seeking a minimum-number-of-sides polygon rather than a minimum-perimeter polygon. Some aspects of the differences and similarities between the two will be discussed later.

The Procedure

(0) Pick an initial point. For example, take the ϵ-expanded point normal to the line segment to the right of (or clockwise from) the upper left vertex of the sampling polygon. This will be the initial current origin of the next steps.

(1) Find the minimum clockwise angle of the vectors from the current origin to the constraint points on ∂J, and the maximum of the angles of the ends of the line segments in $\partial(J + \epsilon)$ for each segment in paired (MPP segment to expanded segment) sequence clockwise. Continue until an end-point of a segment on $\partial(J + \epsilon)$ has an angle greater than the minimum angle of constraint on ∂J, or the initial origin is encountered. If the latter condition holds, close the list of extreme points and exit; otherwise proceed.

(2) Find the intersection of the line from the current origin through the constraint point and the line segment just found, one end of which violates the constraints.

(3) Find the nearest lattice point that gives a line segment to the current origin, which crosses neither ∂J nor $\partial(J + \epsilon)$. Put this point in the extreme point list, and make it the current origin. Return to step 1.

To generalize this procedure to images with concavities, we find intersections of expanded segments of $\partial(J + \epsilon)$ at concave vertices of ∂J. Clearly, extreme points can now be on or close to ∂J as well as $\partial(J + \epsilon)$.

This procedure is repeatable, like that of SCH for the MPP. It is not necessarily convergent to a single limiting polygon; nor is it independent of direction of application (clockwise or counterclockwise), selection of original vertex, or rotation or reflection of the original (or sampled) figures. It should be less efficient in computational complexity and more parsimonious of representation than that of Ramer (1972).

What it does do very well is to allow separate control of sampling and registration error, and of representation or transmission accuracy. It is a natural method for studies of fidelity of area-oriented representations.

An Alternate Procedure

A computationally simpler method for ϵ-approximation may be found by noting that the vertical and horizontal portions of a 4-connected cellular boundary are fairly insensitive to smooth boundary changes. It is the nearly diagonal elements that are very sensitive. One can perform a half-cell expansion, as described in Sklansky (1972c) on the boundary. One can then perform a polygonal approximation similar to the above, using the outer boundary points as extreme points, and the inner convex corners as constraints. This avoids the necessity of solving simultaneous equations, followed by searching for a lattice point that satisfies the constraints. No operations need to be performed on a finer scale than $\epsilon/2$.

Bounds for Sampling and Representation on the Same Net

It is desirable to apply our theory to the intuitively appealing and computationally well-investigated case of both sampling and representation on an ϵ-net or rectangular mosaic. There are (at least) two problems associated with such an application. First, in an extreme case, both ∂J and $\partial(J + \epsilon)$ could force approximation errors of almost (but not quite) $b\epsilon$ on a $b\epsilon$-net. Thus, we have an approximation

$$\epsilon_b \simeq \epsilon(1 - 2b) \tag{21}$$

in some cases. Since ϵ and ϵ_b are nonnegative by assumption, this suggests that the theory might not apply for $b \geqslant 1/2$. Bolour and Cover (1972) used $b = 1/4$, and found $\epsilon_b = \epsilon/2$ in calculating set entropy.

The second problem, more subtle, though well known, is that the uncertainty in boundary location is ϵ for horizontal or vertical segments, but it is smaller for oblique segments, approaching zero for diagonal or $\pm 45°$ segments. From Fig. 9, we see that

$$d = \epsilon(\cos\theta - \sin\theta), \qquad 0 \leqslant \theta \leqslant \pi/4. \tag{22}$$

Similar equations hold for other ranges of θ.

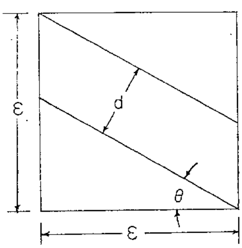

FIG. 9. Illustration of the effect of slope on boundary approximation on a rectangular grid.

We assume uniform probability distributions over both spatial location and angle. We are aware of no assumption that is to be preferred. This assumption gives an average angular-dependent ϵ_a of

$$\epsilon_a = (4/\pi)(2^{1/2} - 1)\epsilon \simeq \epsilon/2. \tag{23}$$

Considering the average over a large number of figures sampled on an ϵ-mosaic, we further assume that the average sampling quantization error is near $\epsilon/2$, and that the truncation error is about 1/2 of the quantization error. Thus, we are led to the inference that an effective ϵ,

$$\epsilon_s \simeq (\epsilon/2)(1 - 1/2)(1/2), \tag{24}$$

may be used either in the perimeter or area bound inequalities to determine the average behavior of the MPP algorithm as an efficient shape descriptor. It should be noted that this average assumes some type of optimum recoding of MPP vertices, which even Sklansky's (1972a) "spine code," for example, does not fully employ. The spine code improves Freeman's chain code (Freeman, 1961) by giving horizontal and vertical run lengths, but this is still more points than the vertices of the MPP require.

Since ϵ enters the bound on the number of extreme points to the $-\frac{1}{2}$ power in both the perimeter and area bounds, Eq. (24) implies that the "average" behavior of an ϵ-net sampling and encoding is at least $2^{3/2}$ less parsimonious than the upper bound on precise ϵ-approximation, where parsimony is measured by number of points needed for representation. Since the extreme example we have found, the circle, is $2^{1/2}$ more parsimonious than this bound, the worst case of precise encoding is probably a factor of 4 more parsimonious than the average for ϵ-sampling and encoding.

Because a factor of 4 is not overpowering, and this section's analysis is imprecise at best (!), the conclusions that we feel should be drawn from this section are the following:

(1) With proper encoding, an ϵ-sampled blob should be encodable to within 1 or $1\frac{1}{2}$ orders of magnitude of the perimeter bound of Bolour and Cover, or of the area bound given here.

(2) With a factor of 4 or 8 oversampling and processing, it should be possible to obtain a representation that is both more accurate and more efficient than sampling at the precision desired.

Conclusions

We have found that convex sets can be ϵ-approximated by polygons with considerably fewer sides than had been previously shown. The number of sides necessary to ϵ-approximate a convex figure is no more than a constant times the fourth root of its area.

It has been shown that concave (nonconvex) figures can be decomposed into convex or nearly convex subfigures. A slight extension of the convex set coding bound applies to the nearly convex figures. It is then shown that any closed figure can be ϵ-approximated by a polygon having a number of sides bounded at most by the area of the figure's convex hull times twice the sum of inflection points, cusps, and maximum winding number. The result appears to be applicable to most continuous curves, with the area defined by closing the curves by one (or at most two) straight line segments.

A procedure for approaching such bounds on ϵ-approximation is given. One important result found is that in order to approach these bounds in practice, it is necessary to sample and to represent the figure on a finer grid, less than one-half the size of the approximation parameter, ϵ. It is found that the effective parameter ϵ_b is substantially smaller than the chosen parameter ϵ unless the sampling-representation grid is considerably smaller than $\epsilon/2$.

Finally, a heuristic argument is given to find the effective ϵ_s of the traditional sampling, representation, and approximation on the same scale. From this analysis, the worst case that we have been able to find for ϵ-approximation with sampling and representation on an $\epsilon/4$-net is four times as efficient as the average behavior of the traditional method on an ϵ-net in a number-of-quantities sense. If the encoded sets are larger than about 4×4 on an ϵ-grid, the same statement holds in a number-of-bits sense.

RECEIVED: April 30, 1975; REVISED: July 1, 1975

REFERENCES

BOLOUR, A., AND COVER, T. M. (1972), "On the Number of Convex Subsets of the Unit Square," Technical Report No. 2 (NSF Grant GK-34363), Department of Statistics, Stanford University, Stanford, Calif.

DUDA, R. O., AND HART, P. E. (1973), "Pattern Classification and Scene Analysis," Wiley-Interscience, New York.

EDEN, M. (1960), A two-dimensional growth process, *in* "Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability," University of California Press, Berkeley, Calif.

EDEN, M. (1968), "On a Procedure for Coding Connected Figures in a Picture," Quarterly Progress Report No. 91, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass.

FREEMAN, H. (1961), On the encoding of arbitrary geometric figures, *I.R.E. Trans. Electronic Comput.* **EC-10**, 260–268.

FREEMAN, H., AND GLASS, J. M. (1969), On the quantization of line-drawing data, *IEEE Trans. Syst. Sci. Cybern.* **SSC-5**, 70–79.

GRAHAM, D. N. (1967), Image transmission by two-dimensional contour coding, *Proc. IEEE* **55**, 336–346.

HUANG, T. S., SCHREIBER, W. F., AND TRETIAK, O. J. (1971), Image processing, *Proc. IEEE* **59**, 1586–1609.

MINSKY, M., AND PAPERT, S. (1972), "Perceptrons: An Introduction to Computational Geometry" (2nd printing), MIT Press, Cambridge, Mass.

MONTANARI, U. (1968), A method of obtaining skeletons using a quasi-Euclidean distance, *J. Ass. Comput. Mach.* **15**, 600–624.

MONTANARI, U. (1969), Continuous skeletons from digitized images, *J. Ass. Comput. Mach.* **16**, 534–549.

MONTANARI, G. U. (1970), On limit properties in digitization schemes, *J. Ass. Comput. Mach.* **17**, 348–360.

PAN, J. W. (1962), "Reduction of Information Redundancy in Pictures," Sc. D. Thesis, MIT, Cambridge, Mass.

RAMER, U. (1972), An iterative procedure for the polygonal approximation of plane curves, *Comput. Graphics Image Process.* **1**, 244–256.

SKLANSKY, J. (1970), Recognition of convex blobs, *Pattern Recog.* **2**, 3–10.

SKLANSKY, J., CHAZIN, R. L., AND HANSEN, B. J. (1972a), Minimum-perimeter polygons of digitized silhouettes, *IEEE Trans. Comp.* **C-21**, 260–268.

SKLANSKY, J., AND NAHIN, P. J. (1972b), A parallel mechanism for describing silhouettes, *IEEE Trans. Comp.* **C-21**, 1233–1239.

SKLANSKY, J. (1972c), Measuring concavity on a rectangular mosaic, *IEEE Trans. Comp.* **C-21**, 1355–1364.

WALPERT, G. (1970), "Image Bandwidth Compression by Contours," Sc. D. Thesis, MIT, Cambridge, Mass.

YOUNG, I. T. (1966), "Television Bandwidth Compression Using Area Properties," S. M. Thesis, MIT, Cambridge, Mass.