MORPHOLOGICAL SKELETON REPRESENTATION AND CODING OF BINARY IMAGES

Petros A. Maragos and Ronald W. Schafer

School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

This paper presents a preliminary study on using Mathematical Morphology to represent and code a binary or a grey-tone image by parts of its skeleton, a thinned version of the image. An image can be uniquely decomposed into skeleton components, and then reconstructed by dilating these components. Since, for a certain category of imagery, the skeleton components possess a lower entropy than the original image, a run-length or entropy coding scheme can be used to achieve representation or transmission of the image at a lower information rate than originally required.

INTRODUCTION TO MATHEMATICAL MORPHOLOGY

Mathematical Morphology, as a method for image analysis, was introduced by Matheron and Serra [1]. Its purpose is the quantitative description of geometrical structures. To extract information from an image object, Morphology "hits" it first with a "structuring element." The interaction with the structuring element transforms the object and reduces it to a sort of caricature which is more expressive than the actual initial phenomenon.

The most fundamental morphological transformations are erosion and dilation: Let X denote a set in the continuous or digital 2-D Euclidean space representing a binary analog or digital image object. Then X^c (complement of X) denotes the image background. Let B be the structuring element, which is another set with a simple geometrical shape, and denote by B_x the translate of B whose center is situated at the point x. Erosion of X by B is the set of all points x such that B_x is included in X (see Fig. 1). Symbolically,

$$X \ominus B=\{x: \; B_x \subset X\} \qquad (1)$$

Dilation of X by B is the set of all points x such that B_x "hits" X; i.e. has a non-empty intersection with X. Symbolically,

$$X \oplus B=\{x: \; B_x \cap X \neq \phi\} \qquad (2)$$

Fig. 1 shows the erosion and dilation of a set X by a disk B. This figure illustrates that erosion is a shrinking operation and dilation is an expanding operation. Erosion and dilation are dual operations w.r. to complementation: Eroding X is equivalent to taking the complement of the dilation of X^c. If we erode X by B and then dilate the set $X \ominus B$ by B, we do not recover X. We reconstitute only a part of X which is simpler and has less details. It may be considered as that part which is most essential morphologically. We call this new set the opening of X w.r. to B:

$$X_B=(X \ominus B) \oplus B \qquad (3)$$

The opening is the domain swept out by all the translates of B which are included in X. This operation smooths the contours of X, cuts the narrow isthmuses, suppresses the small islands and the sharp capes of X.

Although the above operations appear superficially simple, we can perform an enormous variety of image processing and image understanding tasks just by combining erosions and dilations, as is well developed in [1].

SKELETON REPRESENTATION OF BINARY IMAGES

The skeleton is a topologically equivalent thinned version of the image. It can be obtained from morphological transformations which emphasize features of the object associated with its connectivity. In the 2-D continuous space it is defined as follows: Let rB_x denote the disk of radius r centered at the point x. Let $s_r(X)$ denote the set of the centers of the disks rB_x such that: i) rB_x is the maximum disk centered at x and contained in the object X, and ii) the disk rB_x intersects the boundary of X at two or more different places. Then, the skeleton S(X) of X is defined as the set of the centers of the maximum disks inscribable in X, and is a caricature containing information about the shape, size and orientation of X. Some examples of skeletons are shown in Fig. 2. The skeleton S(X) can be obtained from the set union of $s_r(X)$ (Lantuejoul [2]):

$$S(X) = \bigcup_{r>0} s_r(X) = \bigcup_{r>0} \left[(X \ominus rB) / (X \ominus rB)_{drB} \right] \qquad (4)$$

where "U" ("/") represents set union (difference), and dr is the infinitesimal small radius.

Although the skeleton is not a well-digitalizable notion, Serra [1] gives an algorithm for the skeleton of digital binary images sampled on a hexagonal grid.

Our research was focused on three areas: obtaining algorithms for skeletonizing digital binary images on a rectangular grid; using parts of the skeleton to code the image; and extending the above ideas to grey-tone images.

Let R denote the unit-size square of a rectangular grid (see Fig. 3) which is a square of 3×3 pixels, and let nR denote the square R magnified n times which gives a square of (2n+1) × (2n+1) pixels. Then a digital algorithm for S(X) of a rectangularly sampled image object X is

$$S(X) = \bigcup_{n=0}^{n_{max}} s_n(X) = \bigcup_{n=0}^{n_{max}} \left[X \ominus nR) / (X \ominus nR)_R \right] \quad (5)$$

Eq. (5) says that the skeleton subsets $s_n(X)$ form a partition of S(X). Thus, S(X) is obtained by successively eroding X by nR, and then keeping from every eroded set $(X \ominus nR)$ those parts only which consist of angular points and lines without thickness; these parts are the only ones remaining after the set difference between $(X \ominus nR)$ and its opening $(X \ominus nR)_R$. The maximum size n_{max} indicates the square of maximum size after which a further erosion erodes X down to the empty set.

Now, the image X can be exactly reconstructed by dilating the subsets of its skeleton by squares of corresponding size and taking their union:

$$X = \bigcup_{n=0}^{n_{max}} \left[s_n(X) \oplus nR \right] \quad (6)$$

Eqs. (5) and (6) imply that the datum of the image (set X) is equivalent to that of its skeleton S(X) together with the size "n" of the maximum square associated with each point of S(X). In Fig. 4, proceeding from left to right columns, we show an example of an image object X and its erosions $(X \ominus nR)$, the openings of these erosions, the skeleton subsets $s_n(X)$, the dilated subsets, the composition of the skeleton S(X) as the union of the skeleton subsets, and finally the reconstruction of X as the union of the dilated skeleton subsets.

SKELETON IMAGE CODING

According to Shannon's theory of discrete source coding [3] we consider the digitized images as sample functions of a 2-D stochastic process characterized by joint probability distributions of all orders. In practice we measure histograms instead of probability distributions. Consider a 1-D or 2-D block of consecutive pixels $x_1, x_2, \ldots, x_N$, where x_i can be either 1 or 0 according to whether x_i belongs to the image object X or its background X^c respectively. Let $p(x_1, x_2, \ldots, x_N)$ be the Nth-order joint probability of these N pixels. Then the Nth-order joint entropy (in bits/pixel) of the binary image $X \cup X^c$ is defined as

$$H_N(X) = -(1/N) \cdot \sum_{x_1, \ldots, x_N} \cdots \sum p(x_1, \ldots, x_N) \cdot \log_2 p(x_1, \ldots, x_N) \quad (7)$$

As is well known, H_N is a nonincreasing function of N and the limit as $N \to \infty$ is the entropy of the stochastic source. If we consider the 2^N different blocks of N pixels each as our messages, we can employ Huffman coding or other suboptimum coding procedures [4] to achieve transmission rates very close to these Nth-order entropies. Thus, hereafter we will be referring to these Nth-order joint entropies of binary images as their achievable transmission rates.

Since every skeleton subset $s_n(X)$ is a much thinner binary image than X, then its Nth-order entropy, denoted by $H_N(s_n)$, will be much lower than $H_N(X)$. And there might be cases where

$$\sum_{n=0}^{n_{max}} H_N(s_n) << H_N(X) \quad (8)$$

Thus, to transmit $s_n(X)$ we need approximately $H_N(s_n)$ bits/pixel. In addition to the sum of all $H_N(s_n)$ we need information about "n_{max}," which can be taken into account with the trivial amount of $\log_2(M/2)$ bits, for a binary image of M×M pixels.

When (8) holds, we can transmit all the skeleton subsets of X independently at a total rate less than the entropy of the original image, and fully reconstruct X without error as Eq. (6) indicates.

A further reduction in information rate can be achieved by using not all but only some of the skeleton subsets to reconstruct openings (smoothed versions) of the original image:

$$X_{kR} = \bigcup_{n=k}^{n_{max}} \left[s_n(X) \oplus nR \right] \quad (9)$$

That is, if in the union of the skeleton subsets we omit the first k subsets (n=0,1,..,k-1), we reconstruct the opening of X w.r.t. to kR. The larger the k, the fewer subsets we transmit, the more we reduce the information rate, but the smoother is the version X that we reconstruct. As shown in the example of Fig 4, for N=4, the original image X has an entropy of 0.34 bits/pixel. If we use all the skeleton subsets we reconstruct X perfectly at a rate of approximately 0.18 bits/pixel. If we desire to reconstruct only the openings X_R or X_{2R}, we omit the first one or two skeleton subsets and thus we need approximately 0.16 or 0.14 bits/pixel respectively. Table 1 illustrates that more informatively.

TABLE 1
Nth-Order Entropies (bits/pixel) of a Skeleton Reconstructed Image and Its Openings.

N / Image	1	2	4	8
X	0.47	0.22	0.18	0.15
X_R	0.22	0.19	0.16	0.13
X_{2R}	0.20	0.17	0.14	0.10
X_{3R}	0.07	0.06	0.05	0.03

The first-, second-, fourth- and eighth-order entropies of the original binary image without skeleton encoding are 0.79, 0.50, 0.34 and 0.23 respectively. Thus, as shown in Table 1, the sum of the entropies of all or some of the skeleton subsets is smaller than the entropies of the original unencoded image.

SKELETON OF GREY-TONE IMAGES

In grey-tone Morphology [1] the binary erosions and dilations are replaced by "min" and "max" operators respectively. Consider a nonnegative bounded function f(i,j) representing the intensity of a sampled grey-tone image. Let $0 \leq f(i,j) \leq m$ for every integer pair (i,j) in the image support. All the zero-valued image samples will belong by convention to the background of the image object. Erosion or dilation of the function f by the 2-D structuring element R is defined [1] as

$$[f \ominus R](i,j) = \min\left[f(r,s):\ (r,s) \in R_{(i,j)}\right] \qquad (10a)$$

$$[f \oplus r](i,j) = \max\left[f(r,s):\ (r,s) \in R_{(i,j)}\right] \qquad (10b)$$

where $R_{(i,j)}$ denotes the square R centered at the pixel (i,j). The opening f_R of f w.r.t. to R is defined as an erosion followed by dilation.

We provide now a digital algorithm for the skeleton S(f) of f which will be the nonnegative function:

$$S(f) = \sum_{n=0}^{n_{max}} s_n(f) = \sum_{n=0}^{n_{max}} \left[(f \ominus nR) - (f \ominus nR)_R\right] \qquad (11)$$

Eq. (11) is a direct transposition of Eq. (5) where we replaced the binary set union/difference by an algebraic addition/subtraction. Because the opening is an anti-extensive operation ($f_R \leq f$), the skeleton subfunctions in the brackets of (11) will be nonnegative functions. Similarly as in Eq. (6) or (9), the function f or its openings can be reconstructed by summing algebraically all or some of the skeleton subfunctions $s_n(f)$ dilated by nR.

The implications and the coding efficiency of the skeleton of the image function f in terms of entropy considerations are still under investigation.

CONCLUSIONS

The results of this study indicate that a digital binary image can be uniquely decomposed into its skeleton and the maximum inscribable squares, and uniquely reconstructed from its skeleton. The skeleton provides useful information about the shape, size and orientation of an image. For certain categories of images the total entropy of the skeleton subsets is lower than the entropy of the original images. Original 1 bit/pixel test images of irregularly and regularly shaped objects were reconstructed without error by their full skeleton at information rates of ≈0.20 bits/pixel. Smoothed versions of these images required rates of only ≈0.15 bits/pixel. Finally, by using min/max operations instead of binary erosions/dilations, these ideas can be extended to grey-tone images.

REFERENCES

[1] J. Serra, "Image Analysis and Mathematical Morphology," Acad. Press, New York, 1982.

[2] Ch. Lantuejoul," Skeletonization in Quantitative Metallography," in Issues of Digital Image Proc. (R. M. Haralick and J. C. Simon, eds.) series E: Applied Science, No. 34, 1980.

[3] R. G. Gallager, "Information Theory and Reliable Communication," J. Wiley & Sons, Inc., New York, 1968.

[4] T. S. Huang, "Coding of Two-Tone Images," IEEE Trans. on Commun., Vol. COM-25, No. 11, Nov. 1977, pp. 1406-1424.

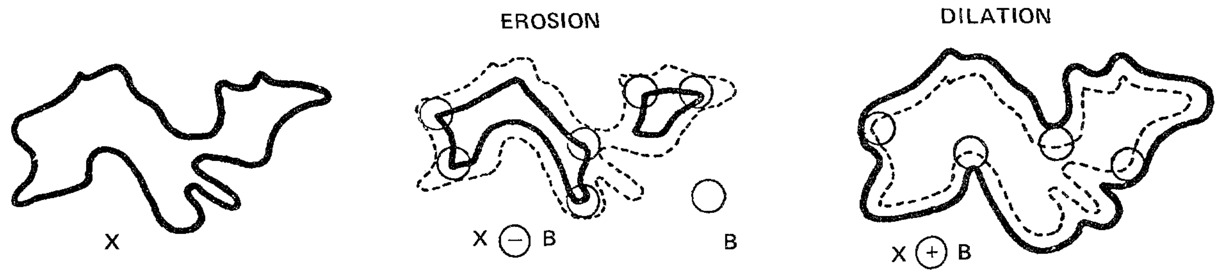

Figure 1 – Erosion and Dilation of a set X by B (after [2]).

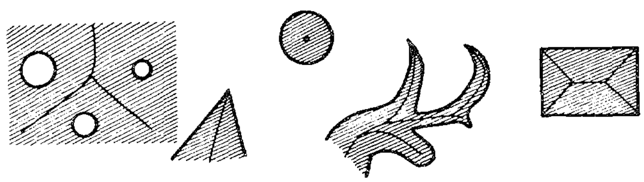

Figure 2 – Examples of Skeletons (after [1]).

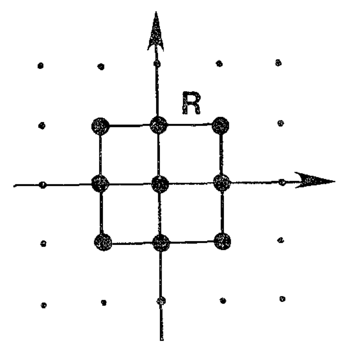

Figure 3 – The 3 x 3 pixels square R on a rectangular grid.

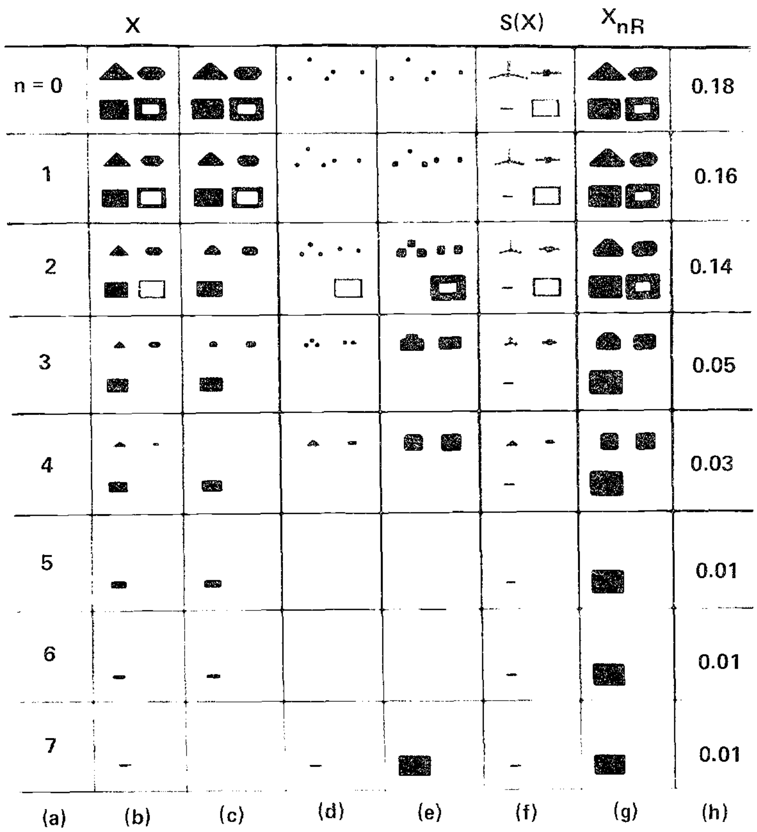

Figure 4 – Step by step decomposition and reconstruction of an image object X by the components of its skeleton S(X) :
(a) size – n – of the structuring square nR
(b) eroded sets ($X \ominus nR$)
(c) openings of the eroded sets ($X \ominus nR$ $)_R$
(d) skeleton subsets $s_n(X)$
(e) dilated skeleton subsets $s_n(X) \oplus nR$
(f) set union of skeleton subsets $s_k(X)$ for k = 7, 6, ... , n-1, n
(g) set union of dilated subsets $s_k(X) \oplus kR$ for k = 7, 6, ... , n-1, n , which gives the opening X_{nR}
(h) sum of the entropies $H_4(s_k)$ of the subsets $s_k(X)$, k = 7, ... , n , which are required to reconstruct the opening X_{nR} of the original object X