MORPHOLOGY-BASED SYMBOLIC IMAGE MODELING, MULTI-SCALE NONLINEAR SMOOTHING, AND PATTERN SPECTRUM

Petros Maragos

Division of Applied Sciences, Harvard University, Cambridge, MA 02138

ABSTRACT.[1] This paper develops a symbolic modeling of images based on their shape-size information. First, multi-scale multi-shape structural distributions in the image are modeled via morphological openings, and a related shape-size descriptor, the pattern spectrum, is developed that can detect critical scales. Then the image is modeled as a nonlinear superposition of simpler parts (the "symbols"), which are translated and scaled shape patterns drawn from a finite collection. The model parameters are found by using the information from openings and pattern spectrum, and via local searches at points of generalized skeletons. All these results appear promising for multi-scale image analysis and shape recognition.

1 Introduction

A main approach in automated recognition [1] of 3-D visual objects or 2-D image silhouettes is based on image decomposition into parts. For example, silhouettes can be modeled as a Boolean combination of smaller binary image objects. Similarly, image depth functions can be modeled as a max-superposition of the 2-D surface functions of simpler 3-D objects. Some related recent research results include [2,3].

These paradigms can be viewed as cases of symbolic image representation and/or description, where the "symbols" are the simpler parts comprising the image and the model involves a nonlinear superposition of modified symbols. This paper develops a symbolic modeling of images based on their "morphology", which is intuitively understood here as the shape-size information contained in the image. Our region-based model attempts to exactly represent an image as a (binary or multilevel) Boolean combination of translated and scaled patterns from a finite collection $\mathcal{K}$, in a *minimal* way agreeing with human intuition. Shape patterns in images may occur at multiple scales [4]-[7]. Hence, we closely relate the model with the development of multi-scale nonlinear filters that smooth the image by probing it with all the patterns of $\mathcal{K}$. A by-product of these multi-shape multi-scale filters is a shape-size descriptor, called pattern spectrum [8,9]. Both the pattern spectrum and the related nonlinear filters are useful tools for multi-scale image analysis. In particular, they aid us significantly in solving the symbolic image modeling problem, which is rigorously formulated based on concepts from mathematical morphology [10,11] and generalized morphological skeletons [16]. Finally, an algorithm is provided that incorporates the above ideas to obtain the model parameters.

[1]This research was supported by a National Science Foundation PYIA under Grant MIPS-86-58150 with matching funds from Bellcore, Xerox and an IBM Departmental Grant, and in part by National Science Foundation under Grant CDR-85-00108.

2 Nonlinear Image Smoothing

By *scale* we define here the *smallest size* of a shape pattern that can fit inside the image. So far [4]-[7], scale has been quantified by varying the average "width" σ of the impulse response (e.g., a Gaussian, local averager) of a linear low-pass filter that smooths the image. This linear filtering approach has gained popularity because of its mathematical tractability, its close relationships with Fourier analysis, and its plausibility for being used at the early stages of the human visual system. However, we also see three rather weak points: 1) *Linear* filters shift and blur important image features such as edges, 2) their implied scale parameter σ is not directly related to our aforementioned size-based definition of scale, and 3) the multi-scale filtered versions of the signal do not correspond to a compact shape representation, except for the obvious one, that is, the difference signals between filtered versions at successive scales. Alternatively, there is a large class of *nonlinear* filters that avoid some or all of the aforementioned problems. They include *median* filters and *opening/closing* [10]-[15] filters. Many such nonlinear (e.g., opening, median) and linear (e.g., averager, Gaussian) smoothing filters can be represented exactly as a minimal superposition of morphological erosions or dilations [8,14,15]. Openings of sets in Euclidean spaces by convex sets of varying size (scale) were introduced by Matheron [10] as an axiomatization of the concept of size. As shown in [17], openings of 1-D boundary curvature functions of continuous binary images by disks do not introduce additional zero-crossings at coarser scales (larger disk radii). A multiresolution approach based on openings was developed in [18]. Median and opening/closing filters can be defined based on a scale parameter and provide signal smoothing by eliminating impulses or narrow peaks/valleys while preserving its edges. In this paper we deal only with openings/closings, because they are directly related to the solution of the symbolic modeling problem and have many other useful properties.

CONTINUOUS SIZE: Let $\mathbf{R}$ and $\mathbf{Z}$ denote, respectively, the set of real numbers and integers. Let B be a compact connected subset of $\mathbf{R}^2$; we call such B a *continuous binary pattern.* If B has size (by convention) one, then the set

$$rB = \{rb : b \in B\} \ , \ r \geq 0, \qquad (1)$$

defines a pattern shaped like B of *size* r. Let $g(x,y)$ be a 2-D real-valued function whose support is a compact connected subset of $\mathbf{R}^2$. Its *umbra* [12] is the set $U(g) = \{(x,y,a) \in \mathbf{R}^3 : g(x,y) \geq a\}$. If we view g as a *graytone pattern* of size one, then we define the function

$$(rg)(x,y) = \sup\{a \in \mathbf{R} : (x,y,a) \in rU(g)\} \qquad (2)$$

as a pattern of continuous size $r \geq 0$. That is, $U(rg) = rU(g)$. The function rg has the same shape as g, but both its support and range are scaled by a factor r.

DISCRETE SIZE: Both (1) and (2) are not useful for defining the size of discrete patterns; e.g., if $B \subseteq \mathbf{Z}^2$ is convex, rB is not convex for $r \geq 2$. This led us to give an alternative definition of discrete size based on set dilation. The *dilation* (also called *Minkowski sum*) of sets X and Y is the set $X \oplus Y = \{a+b : a \in X, b \in Y\} = \bigcup_{p \in Y} X + p$, where $X + p = \{a + p : a \in X\}$ is the *translate* of X by the vector p. Let B be a *discrete binary pattern*, that is, a finite connected subset of the discrete plane $\mathbf{Z}^2$. If B is of size (by convention) one, the sets

$$nB = \underbrace{B \oplus B \oplus \ldots \oplus B}_{n \text{ times}} \qquad (3)$$

define binary patterns of size $n = 0, 1, 2, \ldots$. If B is convex, then nB is shaped like B but has size n. Similarly, let $g(x,y)$ be a *discrete graytone pattern*, that is, a real-valued function whose support is a finite connected subset of $\mathbf{Z}^2$, of size one. If $(f \oplus g)(x,y) = \max_{(i,j)}\{f(x-i, y-j) + g(i,j)\}$ denotes the *dilation* [12] of functions f and g,

$$ng = \underbrace{g \oplus g \oplus \ldots \oplus g}_{n \text{ times}} \qquad (4)$$

defines function patterns of discrete size $n = 0, 1, 2, \ldots$.

Multi-scale Openings and Closings

For sets X, Y in $\mathbf{R}^m$ or $\mathbf{Z}^m$, the *opening* [10] of X by Y is the set $X \circ Y = (X \ominus Y) \oplus Y$, where $X \ominus Y = \{p : Y + p \subseteq X\}$ is the *erosion* of Y from X. Similarly, the *closing* [10] of X by Y is the set $X \bullet Y = (X \oplus Y) \ominus Y$. (For set and function erosion, dilation, opening, and closing, we had used in earlier work [8],[9],[14]-[16] the definitions of [10, 11]; in this paper we use the slightly different definitions of [12,13] because they are simpler.) We henceforth denote a discrete *binary image* by a *set* X in $\mathbf{Z}^2$. Let $B \subseteq \mathbf{Z}^2$ be a fixed pattern. We define

$$X \circ nB = [(\underbrace{X \ominus B) \ominus B \ldots \ominus B}_{n \text{ times}}] \underbrace{\oplus B \oplus B \ldots \oplus B}_{n \text{ times}} \ . \qquad (5)$$

as the *multi-scale opening* of X by B at scale $n = 0, 1, 2, \ldots$. A *dual* multi-scale filter is the closing $X \bullet nB = (X \oplus nB) \ominus nB$. If $n = 0$, then $X \circ nB = X \bullet nB = X$. If B has a regular shape, then $X \circ nB$ and $X \bullet nB$ provide multi-scale nonlinear smoothing of the boundary of X, but they are region-based image operations. The opening suppresses the sharp capes and cuts the narrow (relative to nB) isthmuses of X, because

$$X \circ nB = \bigcup_{nB+z \subseteq X} nB + z \ . \qquad (6)$$

Hence, $X \circ nB$ eliminates from X all objects of size $< n$ (with respect to B), that is, objects inside which nB cannot fit. That is why, we use the size n of nB as synonymous to the scale at which the filter $X \circ nB$ operates. The closing $X \bullet nB$ provides a multi-scale nonlinear smoothing of the background of X by filling in the gulfs and the small (relative to nB) holes of X.

If f is a graytone image function, and g is a graytone pattern on $\mathbf{Z}^2$, we define

$$f \circ ng = (f \ominus ng) \oplus ng = [(\underbrace{f \ominus g) \ominus g \ldots \ominus g}_{n \text{ times}}] \underbrace{\oplus g \oplus g \ldots \oplus g}_{n \text{ times}} \quad (7)$$

as the *multi-scale opening* of f by g at scale $n = 0, 1, 2, \ldots$, where $(f \ominus g)(x,y) = \min_{(i,j)}\{f(x+i, y+j) - g(i,j)\}$ is the *erosion* [12] of f by g. Likewise, the function $f \bullet ng = (f \oplus ng) \ominus ng$ is the *multi-scale closing* of f by g. See Fig. 1 for examples. To implement $f \oplus g$ and $f \ominus g$ we assume that f and g are equal to $-\infty$ outside their supports, where $\mathrm{Spt}(f) = \{(x,y) : f(x,y) \neq -\infty\}$ denotes the *support* of such functions f; see also [11]-[14].

Let g be a binary function equal to 0 inside its support and $-\infty$ elsewhere, and let $B = \mathrm{Spt}(g)$ be its corresponding binary pattern. Then $f \oplus g$ becomes $(f \oplus B)(x,y) = \max_{(i,j) \in B}\{f(x-i, y-j)\}$, and $f \ominus g$ becomes $(f \ominus B)(x,y) = \min_{(i,j) \in B}\{f(x+i, y+j)\}$. These *local min/max* graytone image operations were introduced in [19] and related to binary shrink/expand operations; the latter were used extensively in [20]. The functions $f \circ nB = (f \ominus nB) \oplus nB$ and $f \bullet nB = (f \oplus nB) \ominus nB$ are the multi-scale opening and closing of f by B at scale $n = 0, 1, 2, \ldots$. The nonlinear smoothing effects of the multi-scale opening/closing by B are comparable to the smoothing by g. At large scales n the opening (closing) by nB creates large flat plateaus (sinks) shaped like nB. By contrast, the multi-scale opening (closing) by g will give these summits (sinks) a form of peaks (valleys) shaped like ng; thus, it proceeds much slower than by B, because of the smooth 3-D shaping (e.g., spherical) of g. However, $f \circ nB$ has many computational advantages over $f \circ ng$; see [14]. Further, a nonlinear partial differential equation has been found [21] that models the multi-scale $f \circ rB$ as a dynamic system.

Skeletons and Openings

BINARY IMAGES: Among the various approaches [22] to obtain the medial (symmetric) axis transform, the latter can also be obtained via erosions and openings [11],[16]. Let $X \subseteq \mathbf{Z}^2$ represent a finite discrete binary image and let $B \subseteq \mathbf{Z}^2$ be a binary pattern with $(0,0) \in B$. The *skeleton* *components* of X with respect to B are the sets

$$S_n = (X \ominus nB) \setminus [(X \ominus nB) \circ B] \ , \ n = 0, 1, \ldots, N, \quad (8)$$

where $N = \max\{n : X \ominus nB \neq \emptyset\}$ and $\setminus$ denotes set difference. The S_n are disjoint subsets of X, whose union is the *morphological skeleton* of X and from which we can reconstruct [16] X; i.e.,

$$X \circ kB = \bigcup_{k \le n \le N} S_n \oplus nB \ , \ 0 \le k \le N. \quad (9)$$

Thus, if $k = 0$ (i.e., if we use all the skeleton subsets), $X \circ kB = X$ and we have *exact reconstruction.* If $1 \le k \le N$, we obtain a *partial reconstruction*, i.e., the opening (smoothed version) of X by kB. The larger the k, the larger the degree of smoothing.

GRAYTONE IMAGES: Let $f(x)$, $x \in \mathbf{Z}^2$, be a finite-support discrete graytone image function and let g be a discrete graytone pattern with $g(0) \ge 0$. As the *skeleton components* of f with respect to g we define the nonnegative functions

$$s_n(x) = \begin{cases} f \ominus ng(x), & \text{if } f \ominus ng(x) > (f \ominus ng) \circ g(x) \\ -\infty, & \text{if } f \ominus ng(x) = (f \ominus ng) \circ g(x) \end{cases} \quad (10)$$

where $0 \le n \le N = \max\{n : f \ominus ng \not\equiv -\infty\}$. (For functions $h(x)$, $h \equiv -\infty$ means $h(x) = -\infty \ \forall x \Longleftrightarrow \mathrm{Spt}(h) = \emptyset$.) From the functions s_n we can reconstruct f. Namely, observe that $s_N = f \ominus Ng$ and $f \ominus ng = s_n \vee [(f \ominus ng) \circ g]$, where $(a \vee b)(x) = \max[a(x), b(x)]$. Then, $f \circ kg = [((s_N \oplus g) \vee s_{N-1}) \oplus g \ldots \vee s_k] \oplus kg$. Since function $\oplus$ commutes with $\vee$,

$$f \circ kg = \bigvee_{k \le n \le N} s_n \oplus ng \ , \quad (11)$$

for $0 \le k \le N$. Thus, if $k = 0$, $f \circ kg = f$ and we have *exact reconstruction.* If $1 \le k \le N$, we obtain the opening (smoothed version) of f by kg.

3 Pattern Spectrum

Consider a compact set $X \subseteq \mathbf{R}^2$ and a disk D. If $A(\cdot)$ denotes area, the decreasing function $A(X \circ rD)/A(X)$, $r \ge 0$, was related in [10,11] to probabilistic measures of the size distribution in X. Here we relate these size distributions to a concept of a pattern spectrum for continuous binary images. Then we introduce a pattern spectrum for continuous graytone images and discrete images.

CONTINUOUS IMAGES: We define the *pattern spectrum* of a compact binary image $X \subseteq \mathbf{R}^2$ relative to a *convex* binary pattern $B \subseteq \mathbf{R}^2$ as the nonnegative function

$$PS_X(r, B) = -\frac{dA(X \circ rB)}{dr} \ , \ r \ge 0 \ , \quad (12)$$

and $PS_X(-r, B) = dA(X \bullet rB)/dr$, $r > 0$. The rationale behind our term "pattern spectrum" is the fact that the opening $X \circ rB$ is the union of all $rB + z$ with $rB + z \subseteq X$, that is, of all the patterns shaped like B of size r (located at points z) that can fit inside X. Thus $A(X \circ rB)$ is a measure of the pattern content of X relative to the pattern rB; then (12) measures the differential such pattern content. The *boundary roughness* of X relative to B manifests itself as contributions in the lower-size part of the pattern spectrum; see Fig. 2. Long capes or bulky protruding parts in X that consist of patterns sB show up as isolated impulses (or jumps) in the pattern spectrum around positive $r = s$. Finally, big jumps at negative sizes illustrate the existence of prominent intruding *gulfs* or *holes* in X.

We define the pattern spectrum of a compact-support graytone image f relative to graytone patterns g with convex umbra by the nonnegative function

$$PS_f(r, g) = -\frac{dA(f \circ rg)}{dr} \ , \ r \ge 0 \ , \quad (13)$$

and $PS_f(-r, g) = dA(f \bullet rg)/dr$, $r > 0$, where $A(f)$ is the finite area under the graph of $f(x, y) \ge 0$.

DISCRETE IMAGES: The analog pattern spectrum, except for very simple input signals and very simple probing patterns B or g, is very difficult to compute analytically. To efficiently use arbitrary B or g we extend here the pattern spectrum ideas to discrete binary and graytone images by using definitions (3) and (4) of discrete size. Let $X \subseteq \mathbf{Z}^2$ be a finite binary image and $B \subseteq \mathbf{Z}^2$ a binary pattern. Since $X \circ (n+1)B \subseteq X \circ nB \ \forall n \ge 0$, $A(X \circ nB)$ can only decrease as n increases. Further, $A(X \circ nB) - A[X \circ (n+1)B] = A[X \circ nB \setminus X \circ (n+1)B]$. We define the *pattern spectrum* of X as the nonnegative function

$$PS_X(n, B) = A[X \circ nB \setminus X \circ (n+1)B] \ . \quad (14)$$

Similarly, let $f(x, y) \ge 0$ be a finite-support *graytone image* function on $\mathbf{Z}^2$. We define the *pattern spectrum* of f relative to a discrete graytone pattern g as the nonnegative function

$$PS_f(n, g) = A[f \circ ng - f \circ (n+1)g] \ , \ n \ge 0, \quad (15)$$

where in (15) $A(f) = \sum_{(x,y)} f(x, y)$, and $-$ denotes pointwise function difference. The nonnegativity of (15) follows from $f \circ (n+1)g \le f \circ ng \ \forall n \ge 0$. (The function ordering $a \le b$ means that $a(x) \le b(x) \ \forall x$.) In both (14) and (15) we can extend the pattern spectrum to negative sizes via closings; e.g., $PS_f(-n, g) = A[f \bullet ng - f \bullet (n-1)g]$, $n \ge 1$. Note that $PS_f(n, g) = 0$ for all $n > N = \max\{k : f \ominus kg \not\equiv -\infty\}$ and for all $n < -K$, where K is the minimum size k such that $f \bullet ng = f \bullet kg \ \forall n > k$. Fig. 3 shows a 1-D function f and its pattern spectrum with respect to a binary g whose umbra is a semi-infinite rectangle; i.e., the top 1-D segment of the rectangle is $B = \mathrm{Spt}(g)$. This pattern spectrum of f has impulses at sizes (with respect to B) $n = 1$, 3, and 5, because f contains protruding peaks at these sizes, whereas peaks of size 2 or 4 are not observed. Similarly, f contains many intruding valleys at size 2 which show up as a large impulse in the pattern spectrum at size $n = -2$.

Observe from (9) that $S_n = \emptyset$ implies that $X \circ nB = X \circ (n+1)B$ ($\Longleftrightarrow PS_X(n, B) = 0$). For $1 \le k \le N$,

$$X = X \circ kB \Longleftrightarrow PS_X(n, B) = 0 \quad 0 \le n < k \ . \quad (16)$$

Thus X is smooth to a degree k relative to B (i.e., $X = X \circ kB$) if and only if its first k pattern spectrum samples are zero, or if its first k skeleton components are empty.

Likewise, from (10) it follows that $s_n \equiv -\infty \Longrightarrow PS_f(n,g) = 0$; further, for $1 \le k \le N$,

$$f = f \circ kg \Longleftrightarrow PS_f(n,g) = 0 \quad 0 \le n < k \, . \qquad (17)$$

Thus f is smooth with respect to (i.e., does not contain) peaks of size k relative to g if and only if its first k pattern spectrum samples are zero, or if its first k skeleton components vanish. Likewise for the relationships among the valleys of f, its pattern spectrum values at negative sizes, and the skeleton components of $-f$. Hence, the functions s_n behave as multi-scale shape-size components of f.

Shape-Size Complexity: In the theory [10,11] of random stationary sets $X \subseteq \mathbf{Z}^2$, the size function $\lambda_X(z) = \max\{n : z \in X \circ nB\}$, $z \in X$, can be viewed as a random variable. Its probability function $p_k = \text{Prob}\{\lambda(z) = k\}$ is equal to $PS_X(k,B)/A(X)$. As explained in [9],

$$H(X/B) = -\sum_{n=0}^{N} p_n \log p_n \qquad (18)$$

is the average uncertainty (entropy) of λ. We call it the ***average roughness*** of X relative to B; it quantifies the shape-size complexity of X by measuring its boundary roughness averaged over all depths that B reaches. Thus $H(X/B)$ is maximum ($\log(N+1)$) iff X contains maximal patterns nB at equal area portions in all sizes n, and minimum (0) iff X is the union of maximal patterns of only one size. $H(X/B)$ monotonically decreases as a function of the size k in multi-scale openings of X by kB; i.e., for $0 \le k \le N$,

$$\begin{aligned} H(X \circ kB/B) &= \log A(X \circ kB) - \\ &\tfrac{1}{A(X \circ kB)} \textstyle\sum_{k \le n \le N} PS_X(n,B) \log[PS_X(n,B)] \, . \end{aligned} \qquad (19)$$

All the above ideas can be extended to graytone images f. Thus, the ***average roughness*** of f relative to a graytone pattern g is $H(f/g) = -\sum_{n=0}^{N} q_n \log q_n$, where $q_n = PS_f(n,g)/A(f)$.

4 Symbolic Image Modeling

Henceforth we deal only with *discrete* images. Let $\mathcal{K} = \{B_i : 1 \le i \le M\}$ be a finite collection of M binary patterns (finite 2-D or 1-D subsets of $\mathbf{Z}^2$, preferrably convex), which all have size one and contain the origin of $\mathbf{Z}^2$. In gestalt psychology, the *law of simplicity (Pragnanz)* states that ***every stimulus pattern is seen in such a way that the resulting structure is as simple as possible.*** Attempting to quantify this law, we model a finite binary image $X \subseteq \mathbf{Z}^2$, as a *minimal union of maximal patterns* from $\mathcal{K}$; i.e.,

$$X = \bigcup_{i=1}^{I} (n_i B_{\omega_i}) + p_i \, , \qquad (20)$$

where $1 \le \omega_i \le M$, and $(n_i B_{\omega_i}) + p_i$ is a pattern of size $n_i \ge 0$, located at p_i, shaped similarly to some $B_{\omega_i} \in \mathcal{K}$. By "minimal union" we mean that the total number, I, of patterns required to cover X exactly should be minimum over $\mathcal{K}$. By "maximal patterns" we mean that they are not included (by set inclusion) in any other pattern of the same shape and larger size or of different shape; see Fig. 4.

To model a discrete ***graytone*** image function $f(x)$, $x \in \mathbf{Z}^2$, assume first a finite collection $\mathcal{G} = \{g_i : 1 \le i \le M\}$ of M discrete graytone patterns (binary or multilevel, 1-D or 2-D). Then we model f as a ***minimal*** (with respect to the number I required) ***max-superposition of maximal*** (with respect to the function ordering $\le$) ***patterns*** from $\mathcal{G}$:

$$f(x) = \max_{1 \le i \le I} \{n_i g_{\omega_i}(x - y_i) + c_i\} \, , \qquad (21)$$

where $x, y_i \in \mathbf{Z}^2$ and $c_i \in \mathbf{R}$. Observe that (21) is equivalent to modeling $U(f)$ in $\mathbf{Z}^2 \times \mathbf{R}$ as a minimal union of maximal set patterns $U(g)$, $g \in \mathcal{G}$; further, shifting $U(g)$ by $(y,c) \in \mathbf{Z}^2 \times \mathbf{R}$ is equivalent to shifting the argument of $g(x)$ by y and its amplitude by c.

We consider models (20) and (21) as symbolic image representations, because we view each pattern in $\mathcal{G}$ as a separate symbol corresponding to a set of numerical values. Since we require maximal patterns, this symbolic model represents the image beyond (more coarsely than) its numerical values. If the sizes n_i are all 0, then all the scaled patterns reduce to single points or numbers, and we reobtain the original numerical representation of the image. The modeling problem consists of finding the I triples (ω_i, n_i, p_i); in (21) $p_i = (y_i, c_i)$.

4.1 Binary Images

We can rewrite (20) as

$$X = \bigcup_{i=1}^{M} \bigcup_{n=0}^{N_i} H_{n,i} \oplus nB_i \, , \qquad (22)$$

where $N_i = \max\{n : X \circ nB_i \ne \emptyset\}$, and $H_{n,i}$ denotes the (perhaps empty) set of locations of maximal patterns nB_i. Note that $I = \sum_n \sum_i A(H_{n,i})$. Let $S_{n,i}$ denote henceforth the n-th skeleton component (given by (8)) of X with respect to B_i. If $p \in H_{n,i}$, then $(nB_i) + p$ is maximal in X, which implies [16] that $p \in S_{n,i}$. Hence, for $0 \le n \le N_i$, $1 \le i \le M$,

$$H_{n,i} \subseteq S_{n,i} \subseteq X \ominus nB_i \, . \qquad (23)$$

Thus the solution of the modeling problem is very closely related to morphological skeletonization. Let $N = \max_i\{N_i\}$. Then the problem consists of finding the $(N+1) \times M$ array of set-valued entries $H_{n,i}$ under the two conditions of (23) and the minization of I. We solve this problem by breaking it down into two parts (A and B) and by exploiting the results (9) and (23):

Problem A: We decide first which type of patterns at which scales (sizes) are present. That is, we find a ***shape-size containment array*** $C(n,i)$, which is equal to 1 if $H_{n,i} \ne \emptyset$ and 0 if $H_{n,i} = \emptyset$. This is done by covering X with a collage of the sets

$$D_{n,i} = (X \circ nB_i) \setminus [X \circ (n+1)B_i] \, , \; 0 \le n \le N_i, \qquad (24)$$

for all patterns B_i and by finding which $D_{n,i}$ are needed for a minimal covering of X. (Obviously, $PS_X(n, B_i) = 0 \Longrightarrow$

$D_{n,i} = \emptyset \Longrightarrow C(n,i) = 0.$)

Problem B: For each pattern B_i and for each size n with $C(n,i) = 1$, we search the skeleton component $S_{n,i}$ for a minimal subset $H_{n,i}$ that (together with the maximal patterns already established) guarantees the exact reconstruction of the smoothed version $X \circ nB_i$.

We solve problems A and B as follows. Let $\chi_{n,i}$ denote the ***characteristic function*** of the set $D_{n,i}$; i.e., $\chi_{n,i}(z) = 1$ if $z \in D_{n,i}$ and $\chi_{n,i}(z) = 0$ if $z \notin D_{n,i}$.

Algorithm A: We create a pseudo-graytone function $\alpha(r,s)$ by summing the functions $\chi_{n,i}$ for all (n,i). Then initially α is equal to M times the characteristic function of X. Given a certain scanning order among all (n,i) we eliminate (the protrusion, and hence the contribution, of) a pattern B_j at size k by subtracting from α the function $\chi_{k,j}$, provided that $D_{k,j} \neq \emptyset$ and the modified α-function is ≥ 1 at all $(r,s) \in X$. We repeat the same procedure for all possible scanning orders and choose the scanning that yields a modified α-function with the minimum area; this is equivalent to ***searching for a minimal redundancy in covering*** X. Then we set $C(n,i)$ equal to 1 iff $D_{n,i}$ has not been eliminated during the optimum scanning, and 0 otherwise. To obtain the ***optimum solution*** we must examine all possible scannings whose total number is $T = (\sum_{i=1}^{M} N_i)!$. This makes it computationally formidable since its complexity grows as $(M \cdot N)!$, where N is proportional to the image diameter. Henceforth we adopt a ***suboptimum*** solution to problem A. First, we find the pattern spectrum of X relative to all patterns in $\mathcal{K}$. We also compute the ***multi-scale average roughness array*** $R(n,i) = H(X \circ nB_i / B_i)$ given by (19); see Fig. 5 for an example. Then in trying to eliminate the function $\chi_{n,i}$ from the α-function, we choose a ***single scanning*** determined as follows. We scan sizes at ascending order proceeding from $n = 0$ to $n = N$, because substructures at small sizes are more probable to be covered by protrusions at larger sizes. For each size n, we scan the pattern indices i in an order that corresponds to decreasing average roughness $R(n,i)$. Namely, we try to eliminate first the pattern nB_j if $R(n,j) = \max_i\{R(n,i)\}$, because it is the least likely to contribute to X large protrusions at scale n; then we try the pattern with the next largest $R(n,i)$. Fig. 5 shows the array $C(n,i)$ that resulted from applying the above algorithm to the image of Fig. 5a.

Algorithm B: The solution of problem A yields $C(n,i)$ and a modified α-function, call it β. We exploit (23) and (9), which impies that to go from $X \circ (n+1)B_i$ to $X \circ nB_i$ we need only the information in $S_{n,i}$. Thus, first we skeletonize X with respect to all B_i for which $C(n,i) = 1$ for at least one n. Second, we adopt a scanning order of $C(n,i)$ using $R(n,i)$, as in the solution of problem A. Following this scanning, for each (n,i) with $C(n,i) = 1$, we subtract from β the characteristic function of $D_{n,i}$ and then add the characteristic functions of all $(nB_i) + p$ where p spans $S_{n,i}$. Then, scanning $S_{n,i}$ in a certain order, we eliminate a point p iff subtracting the characteristic function of $(nB_i) + p$ from β leaves $\beta(r,s)$ with values ≥ 1 $\forall (r,s) \in D_{n,i}$. The set $H_{n,i}$ consists of the points in $S_{n,i}$ that were *not* eliminated during the above procedure. An optimum solution would involve trying all $[A(S_{n,i})]!$ scannings of $S_{n,i}$ and choose the one that gives the smallest $A(H_{n,i})$. After having used the suboptimum algorithm A, we applied algorithm B (with a *single scanning* of each $S_{n,i}$) to the image X of Fig. 5 and found that X can be modeled exactly as a union of five maximal patterns: a line at 135° of size 5, a triangle of size 7, and three squares of size 1, 3, and 5. Their locations are shown with ● in Fig. 5a; note that the locations of the 135°-line and the size-3 square coincide.

4.2 Graytone Images

1-D IMAGES: Let $f(x)$, $x \in \mathbf{Z}$, be a finite-support nonnegative function with ***discrete amplitude range*** (e.g., $f(x) \in \mathbf{Z}$). We can rewrite (21) as

$$f(x) = \max_{i=1}^{M} \max_{n=0}^{N_i} \{(h_{n,i} \oplus ng_i)(x)\}\ , \tag{25}$$

where all the $g_i \in \mathcal{G}$ are 1-D functions, $N_i = \max\{n : f \circ ng_i \not\equiv -\infty\}$, and the functions $h_{n,i}(x)$ contain isolated (morphological) impulses. That is, at each $y \in \mathbf{Z}$, $h_{n,i}(y) = c \in \mathbf{Z}$ iff the shifted and scaled pattern $[ng_i(x-y)] + c$ is maximal in f; otherwise, $h_{n,i}(y) \equiv -\infty$. See Fig. 6 for an example. The problem is to find all the $h_{n,i}$, some of which may be empty (be everywhere $-\infty$).

The key idea in our solution of (25) is that symbolically modeling f is equivalent to symbolically modeling its umbra. Namely, we *modify* the umbra of f as $U(f) = \{(x,z) : 0 \leq z \leq f(x)\} \subseteq \mathbf{Z}^2$ to include only the points above the x-axis and we view $U(f)$ as a finite discrete binary image. Then we solve (25) by following algotithms A and B of section 4.1 after making the following adjustments: For each size n and pattern g_i, let $\psi_{n,i}(x,z)$ denote the characteristic function of the set $[U(f \circ ng_i) \setminus U(f \circ (n+1)g_i)] \subseteq U(f)$. Form the α-function by adding all $\psi_{n,i}$, which take values in a y-dimension (see Fig. 6). Form a roughness array $R(n,i)$ from the pattern spectrum of f relative to all g_i. Then obtain the shape-size containment array $C(n,i)$ using algorithm A. Next obtain the n-th skeleton components $s_{n,i}$ (given by (10) of f with respect to all g_i for which $C(n,i) = 1$ for at least one n. The β-function of algorithm B is formed for each (n,i) with $C(n,i) = 1$ by adding all $\psi_{k,j}$ with $C(k,j) = 1$ and $(k,j) \neq (n,i)$, as well as adding the contributions from all points in $s_{n,i}$, i.e., the characteristic functions of the (modified) umbrae of $ng_i(x - x_0) + c_0$ whenever $s_{n,i}(x - x_0) = c_0 \in \mathbf{Z}$. Then the algorithm B tries to eliminate, if possible, points x_0 from the support of $s_{n,i}$ and thus yields the functions $h_{n,i}$.

2-D IMAGES: For a 2-D image function $f(x,y)$ with discrete suppport and range, the solution of the modeling problem (25) is a very straightforward extension of the above solution for 1-D image functions. Namely, the graytone patterns g_i can be 2-D or 1-D, and the modified umbrae of f and ng_i are finite subsets of the 3-D space $\mathbf{Z}^3$. In addition, the characteristic functions of the differences among umbrae of successive openings of f and of the umbrae of ng_i will be functions with arguments (x,y,z) and values in a 4-th dimension. The rest of the procedure remains the same.

5 Discussion

Openings and closings are useful nonlinear filters that can complement linear filters in multi-scale image analysis. Their smoothing is equivalent to eliminating skeleton components and reconstructing the image from its pruned skeleton. Further they provide the pattern spectrum, whose locations of large jumps indicate the existence of protruding or intruding substructures in the image at those scales. A solution to the symbolic modeling problem was obtained by using the pattern spectrum to guide us in scanning the shape-size domain, and by restricting (without losing any optimality) the searching for the pattern locations only inside the sparce supports of morphological skeletons. We have experimentally observed that this (generally suboptimum) solution provides the optimum answer (both mathematically and in good agreement with human intuition) for simple small synthetic images and small pattern collections. Our solution provides an *exact* image representation. To cope with *noise*, certain robustness is inherently built in our formulation that searches for minimal subsets of skeletons. In addition, the image can be pre-smoothed via openings/closings. If the image contains 1-D structures that risk being eliminated by smoothing with a 2-D pattern, a max-superposition of openings by 1-D patterns at various orientations would both provide smoothing and preserve the 1-D structures. To use *rotated* patterns in (22), we must equip the pattern collection with some rotated versions of the same pattern. However, rotation of patterns is not needed for modeling 1-D multilevel functions.

References

[1] P.J. Besl and R.C. Jain, "Three-Dimensional Object Recognition," *Comp. Surv. ACM*, 17, Mar. 1985.

[2] A.P. Pentland, "Recognition by Parts," in *Proc. 1st ICCV*, London, 1987.

[3] J. G. Verly, B. D. Williams and D. E. Dudgeon, "Automatic Object Recognition from Range Imagery using Appearance Models," *Proc. Worksh. Comp. Vision*, Miami, Nov. 1987.

[4] D. Marr and E. Hildreth, "Theory of Edge Detection," *Proc. R. Soc. Lond. B 207*, pp.187-217, 1980.

[5] A. Witkin, "Scale-space Filtering," in *Proc. IJCAI*, Carlsrue, W. Germany, 1983.

[6] A. Yuille and T. Poggio, "Scaling Theorems for Zero Crossings," *IEEE Trans. PAMI-8*, Jan. 1986.

[7] P. J. Burt and E.H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. COM-31*, Apr. 1983.

[8] P. Maragos, "A Unified Theory of Translation-Invariant Systems With Applications to Morphological Analysis and Coding of Images", Ph.D. dissertation, Georgia Inst. Technology, Atlanta, GA, July 1985.

[9] ——, "Pattern Spectrum of Images and Morphological Shape-Size Complexity," in *Proc. IEEE ICASSP'87*, Dallas,TX, Apr. 1987.

[10] G. Matheron, *Random Sets and Integral Geometry*, Wiley, 1975.

[11] J. Serra, *Image Analysis and Mathematical Morphology*, Acad. Press, 1982.

[12] S. R. Sternberg, "Grayscale Morphology," *CVGIP 35*, p.333, 1986.

[13] R.M. Haralick, S.R. Sternberg, and X. Zhuang, "Image Analysis Using Mathematical Morphology," *IEEE Trans. PAMI-9*, July 1987.

[14] P. Maragos and R. W. Schafer, "Morphological Filters - Part I: Their Set-Theoretic Analysis and Relations to Linear Shift-Invariant Filters," *IEEE Trans. ASSP-35*, Aug. 1987.

[15] ——, "Morphological Filters - Part II: Their Relations to Median, Order-Statistic, and Stack Filters," *IEEE Trans. ASSP-35*, Aug. 1987.

[16] ——, "Morphological Skeleton Representation and Coding of Binary Images", *IEEE Trans. ASSP-34*, Oct. 1986.

[17] M. Chen and P. Yan, "A Multiscaling Approach Based on Morphological Filtering," submitted to *IEEE Trans. PAMI.*

[18] R. M. Haralick, C. Lin, J.S.J. Lee and X. Zhuang, "Multiresolution Morphology," in *Proc. 1st ICCV*, London, 1987.

[19] Y. Nakagawa and A. Rosenfeld, "A Note on the Use of Local Min and Max Operations in Digital Picture Processing", *IEEE Trans. SMC-8*, Aug.1978.

[20] K. Preston, Jr., M. J. B. Duff, S. Levialdi, P. E. Norgren, and J-I. Toriwaki, "Basics of Cellular Logic with Some Applications in Medical Image Processing", *Proc. IEEE*, 67, May 1979.

[21] R. W. Brockett, Lectures in Nonlinear Science, UC Berkeley Summer School, 1987.

[22] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, Acad. Press, 1982.

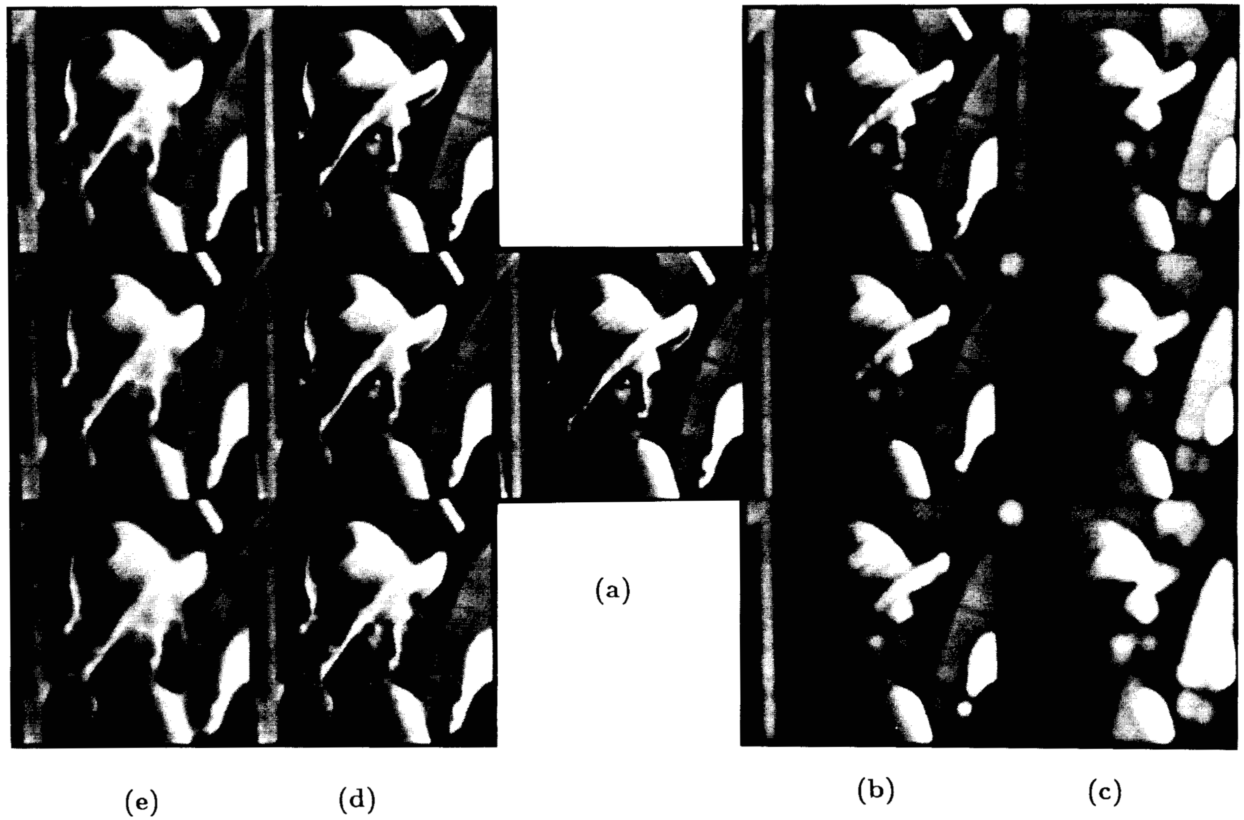


Figure 1. Multi-scale openings and closings. (a) Graytone image f (256×256 pixels). (b) $f \circ ng$, $n = 1, 2, 3$ (top to bottom); the discrete pattern g is defined by $g(x, y) = 5\sqrt{5 - x^2 - y^2}$ if $0 \leq x^2 + y^2 \leq 5$, and $g(x, y) = -\infty$ if $x^2 + y^2 > 5$. (c) $f \circ ng$, $n = 4, 5, 6$ (top to bottom). (d) $f \bullet ng$, $n = 1, 2, 3$ (top to bottom). (e) $f \bullet ng$, $n = 4, 5, 6$ (top to bottom).

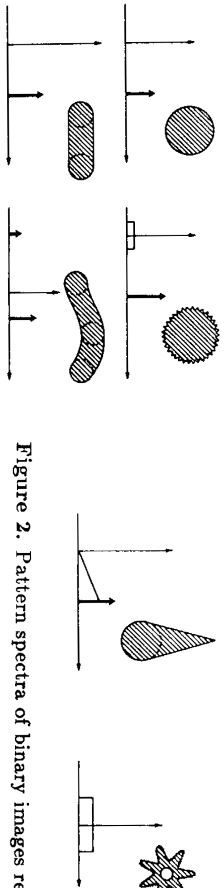


Figure 2. Pattern spectra of binary images relative to a disk.

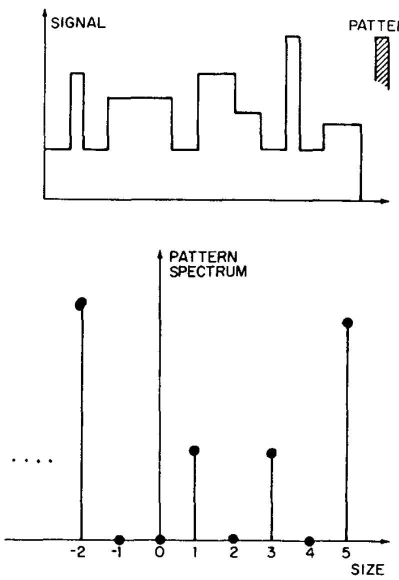


Figure 3. Pattern spectrum of a 1-D multilevel signal.

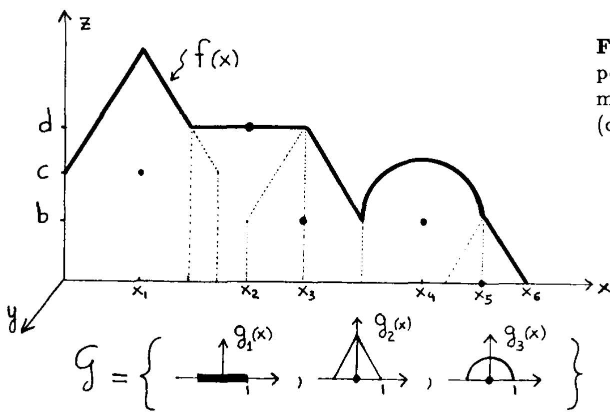
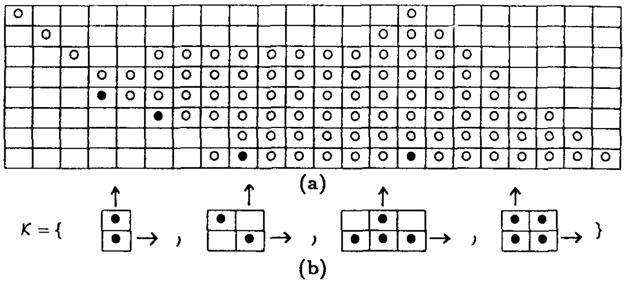


Figure 6. A function $f(x)$ modeled minimally as a max-superposition of 5 maximal patterns: $rg_2(x-x_1)+c$, $sg_2(x-x_3)+b$, $tg_2(x-x_5)$, $vg_1(x-x_2)+d$, $wg_3(x-x_4)+b$, where $r=x_1$, $s=v=x_3-x_2$, $t=x_6-x_5$, and $w=x_4-x_3-s$. (Solid line shows graph of f; dotted line shows occluded graphs of maximal patterns.)



(c)

n \ i				
0	5	3	4	6
1	6	0	0	6
2	3	3	9	6
3	16	0	16	16
4	5	5	0	5
5	36	60	0	54
6	14	14	0	0
7	8	8	64	0

(d)

n \ i				
0	2.55	1.67	1.33	1.88
1	2.38	1.51	1.12	1.65
2	2.17	1.51	1.12	1.38
3	2.01	1.35	0.72	1.08
4	1.61	1.35	0	0.42
5	1.32	1.09	0	0
6	0.95	0.95	0	0
7	0	0	0	0

(e)

n \ i				
0	0	0	0	0
1	0	0	0	1
2	0	0	0	0
3	0	0	0	1
4	0	0	0	0
5	0	1	0	1
6	0	0	0	0
7	0	0	1	0

Figure 5. (a) A discrete binary image X. (○ and ● are points of X; points with ● show the locations-origins of maximal patterns.) (b) $\mathcal{K}=\{B_i : 1 \leq i \leq 4\}$.
(c) $PS_X(n, B_i)$. (d) $R(n,i)$. (e) $C(n,i)$.

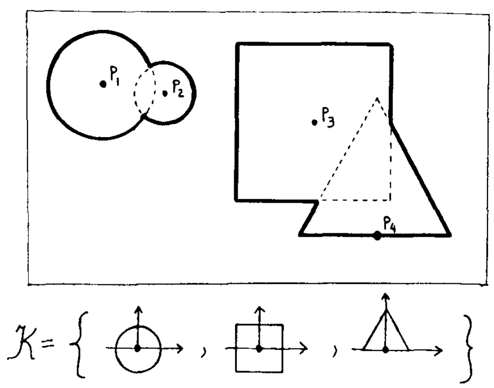


Figure 4. Binary image X modeled as the union of $I=4$ maximal patterns located at p_i. (Solid line shows boundary of X; dotted line shows occluded boundaries of maximal patterns.)