chapter 3

Scale space filtering: a new approach to multi-scale description*

Andrew P. Witkin
Fairchild Laboratory for Artificial Intelligence Research

QUALITATIVE DESCRIPTION AND THE PROBLEM OF SCALE

The waveform depicted in Figure 1 typifies a broad class of intricate, ill-behaved physical signals—this one happens to be an intensity profile from a natural image. In what primitive terms shall we talk about, reason about, learn about, and interpret signals like this one? It is universally agreed that the raw numerical signal values are the *wrong* terms to use for tasks of any sophistication. Plainly, our initial description ought to be as compact as possible, and its elements should correspond closely to meaningful objects or events in the signal-forming process. Unstructured numerical descriptions are inadequate on both counts.

In the case of one-dimensional signals, local extrema in the signal and its derivatives—and intervals bounded by extrema—have frequently proven to be useful descriptiive primitives: although local and closely tied to the signal data, these events often have direct semantic interpretations, for example, as edges in images. A description that characterizes a signal by its extrema and those of its first few derivatives is a *qualitative* description of exactly the kind we were taught to use in elementary calculus to "sketch" a function (Figure 2). Analogous topographic features for two-dimensional functions have been catalogued in Haralick, Watson, and Laffey (1982). Although this paper treats the one-dimensional case only, an extension of the methods presented here to two dimensions is under development.

*I thank Richard Duda and Peter Hart, as well as J. Babaud of Schlumberger, Ltd., for their help and encouragement. Marty Tenenbaum, Harry Barrow, and John Mohammed provided useful comments on earlier versions of this paper.

Figure 1 A typically complicated natural waveform (an intensity profile drawn from an image).

Figure 2 A function characterized by its extrema and those of its first derivative.

A great deal of effort has been expended to obtain this kind of primitive qualitative description, both in one dimension and two (for overviews of this literature, see Ballard & Brown, 1982; Pavlidis, 1977; Rosenfeld & Kak, 1976) and the problem has proved extremely difficult. The problem of *scale* has emerged consistently as a fundamental source of difficulty. Any non-trivial local measurement–including a derivative operator–has to depend on the value of the signal at two or more points, situated on some neighborhood around the nominal point of measurement. The measurement thus depends not only on the signal itself but on the spatial extent of this neighborhood, that is, on a *parameter of scale*. Since the scale parameter influences the measured derivatives, it also influences the qualitative description obtained from the derivatives' extrema. In short, we cannot get any description without specifying the scale of measurement, and different scales yield different descriptions.

How shall we decide which scale, and therefore which description, is correct? The answer, it appears, is that no single scale is categorically correct: the physical processes that generate signals such as images act at a variety of scales, and none is intrinsically more interesting or important than another. The problem is less to distinguish meaningful events from meaningless noise, than to distinguish one process from another, and to organize what we see in a manner that as nearly as possible reflects the physical and causal structure of the world. This and similar lines of thinking have sparked considerable interest in the scale problem and in multi-scale descriptions (see Ballard & Brown, 1982; Hong, Shneier, & Rosenfeld, 1981; Marr & Hildreth, 1979; Marr & Poggio, 1979; Rosenfeld & Thurston, 1971).

Related Work

A radical view of the scale problem is put forward by Mandelbrot (1977), who argues that many physical processes and structures are best modeled by a class of non-differentiable functions called *fractals*. Mandelbrot takes the stance that familiar notions such as length, surface area, slope, and surface orientation, ought to be abandoned entirely in favor of global measures of the processes' behavior over scale (e.g., rate of change of measured arc length.) Fractal models often capture striking regularities in seemingly chaotic processes, as evidenced by their utility in computer graphics. Pentland (1983) has recently argued convincingly for their usefulness in perception as well. However, fractal models cannot supplant the more usual representation of shape: to the mountain climber, knowing what *kind* of mountain he's on is of little help. He needs to know where next to place his hands and feet.

Marr (1982) argued that physical processes act at their own intrinsic scales–for example, a regular patchwork of wheatfields, a stalk of wheat, and the grains on a stalk of wheat–and that each should be described separately. He described the image by the zero-crossings in its convolution with the laplacian of a gaussian, at several fixed scales, apparently in the expectation that these channels would cleanly separate physically distinct processes. Unfortunately, in our experience, they generally fail to do so. Indeed, given that the channels' scales are fixed globally and without regard to the structure of the data, it would be astonishing if they succeeded. One is left with a collection of apparently unrelated descriptions, none of which is quite right. The prospect of deciding which description to use when–or worse still, of picking and choosing among the various channels to build meaningful descriptions–is unappealing. Additionally, Marr proposed the "coincidence assumption" stating that only features that spatially coincide at all scales are physically significant, although no justification was offered for this idea.

Hoffman (forthcoming) recently undertook to replace the fixed-scale sequence of descriptions by descriptions at one or more "natural" scales: measuring the tangent to a curve as a function of scale, he selected as natural scales those points at which the tangent varied least with respect to scale. This approach may offer a significant improvement over fixed channels.

The Scale-Space Approach

Scale-space filtering addresses two distinct problems (drawing a distinction that Marr and others failed to draw clearly). First, we observe different extrema at different scales, any of which might prove meaningful. How then shall we fully characterize the extrema over a broad range of scales? Must we find and describe them independently in each channel, as did Marr, or is it possible to construct a more organized and unified description?

Second, how shall these primitives, whatever they turn out to be, be grouped or *organized* to best reflect the organization of the generating process? Scale is just one among many bases for grouping–along with symmetries, repetitions, and so forth. Where processes really are cleanly separated by scale, we might hope to use scale to group their constituent events; however, scale cannot do the whole job. (For an extended discussion of perceptual organization, see Witkin, 1983; Witkin & Tenenbaum, 1983).

The scale-space description. Our solution to the first problem–that of building a unified description encompassing a broad range of scales–begins with the observation that the descriptions we obtain at nearby scales appear to have a great deal in common: we often see extrema in nearly the same locations, although we are liable to see a number of extrema in the finer channel with no counterparts in the coarser one. (Presumably the same observation motivated Marr's "coincidence assumption.") Is there any well-defined sense in which extrema observed at different scales can be said to correspond, to manifest the same event seen through different filters rather than two unrelated events? We will answer this question by treating scale as a continuous parameter, considering the effect of a very small scale change in scale. (We take as a scale parameter the standard deviation, σ, of a gaussian convolved with the signal.) As we decrease σ continuously, beginning at a coarse scale, we observe two distinct effects on the extrema: (a) existing extrema move continuously along the signal axis, and (b) new extrema occasionally appear at singular points. These effects are best visualized in terms of the surface swept out on the (x,σ) plane by varying σ. We call the (x,σ) plane *scale space*, and the surface, the *scale-space image* of the signal. The extrema, viewed on the scale-space image, form contours. The tops of these contours are the singularities, the points above which a continuously moving extremum vanishes (look ahead to Figure 6).

On the assumption that each of these contours, in general, reflects a single physical event, we take the contour, rather than its constituent points, as the unit of description. The *scale* of the event is the scale at which the contour vanishes, and its *location* in the signal domain is its location at the finest observable scale. The raw qualitative description thus produced consists of a single sequence of extrema, each observable on a definite range of scales, and each referred to a definite signal-domain location. This unified representation exhaustively describes the qualitative structure of the signal over all observed scales. Each element of the description has a scale associated with it (the vanishing scale) but the description itself is not partitioned into arbitrary channels or levels. The description makes the scale of each event available as a basis for organization, but does not prematurely impose a rigid organization.

Organization and scale. We next consider the role of scale as a means of organizing the raw description. A scale-structured representation, called the interval tree, is introduced. Building on the raw scale-space description, and in particular on the singular points at which new extrema appear, the interval tree captures the coarse-to-fine unfolding of finer and finer detail. The tree is used to

generate a set of descriptions, varying the scale in local, discrete steps that reflect the qualitative structure of the signal. Although highly constrained, this family of descriptions appears to capture perceptually salient organizations.

Additionally, a stability criterion is applied to produce "top-level sketches" of signals, as a starting point for matching and interpretation.

THE SCALE-SPACE IMAGE: DEFINITION

We must begin by introducing a parameter of scale, to define what we mean, for example, by "the slope at point x and scale σ." We have chosen to convolve the signal with a gaussian, taking the gaussian's standard deviation as the scale parameter. Although scale-dependent descriptions may be computed in many ways, the gaussian convolution is attractive for a number of its properties, amounting to "well-behavedness": the gaussian is symmetric and strictly decreasing about the mean, and therefore the weighting assigned to signal values decreases smoothly with distance. The gaussian convolution behaves well near the limits of the scale parameter, σ, approaching the unsmoothed signal for small σ, and approaching the signal's mean for large σ. The gaussian is also readily differentiated and integrated.

While the gaussian is not the only convolution kernel that meets these criteria, a more specific motivation for our choice is a property of the gaussian convolution's zero-crossings (and those of its derivatives): as σ decreases, additional zeroes may appear, but existing ones cannot, in general, disappear; moreover, of convolution kernels satisfying "well behavedness" criteria (roughly those enumerated above,) the gaussian is the *only* one guaranteed to satisfy this condition (Babaud, Witkin, & Duda, 1983). This is an important property because it means that all the extrema observed at any scale are observable at the finest scale, which, as we shall see, greatly simplifies the description.

The gaussian convolution of a signal $f(x)$ depends both on x, the signal's independent variable, and on σ, the gaussian's standard deviation. The convolution is given by

$$F(x,\sigma) = f(x) * g(x,\sigma) = \int_{-\infty}^{\infty} f(u) \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-u)}{2\sigma^2}} \, \mathrm{d}u, \tag{1}$$

where "$*$" denotes convolution with respect to x.

F defines a surface on the (x,σ) plane, the surface swept out as the gaussian's standard deviation is smoothly varied. We call the (x,σ)-plane *scale space*, and the surface, F, defined in (1), the *scale-space image* of f.[1] Figure 3 graphs a se-

[1] It is actually convenitent to treat log σ as the scale parameter, as uniform expansion or contraction of the signal in the x-direction will cause a translation of the scale-space image along the log σ axis. All illustrations portray σ on a log scale.

quence of gaussian smoothings at increasing σ, which are constant-σ profiles from the scale-space image. Figure 4 portrays the scale-space image as a surface in perspective.

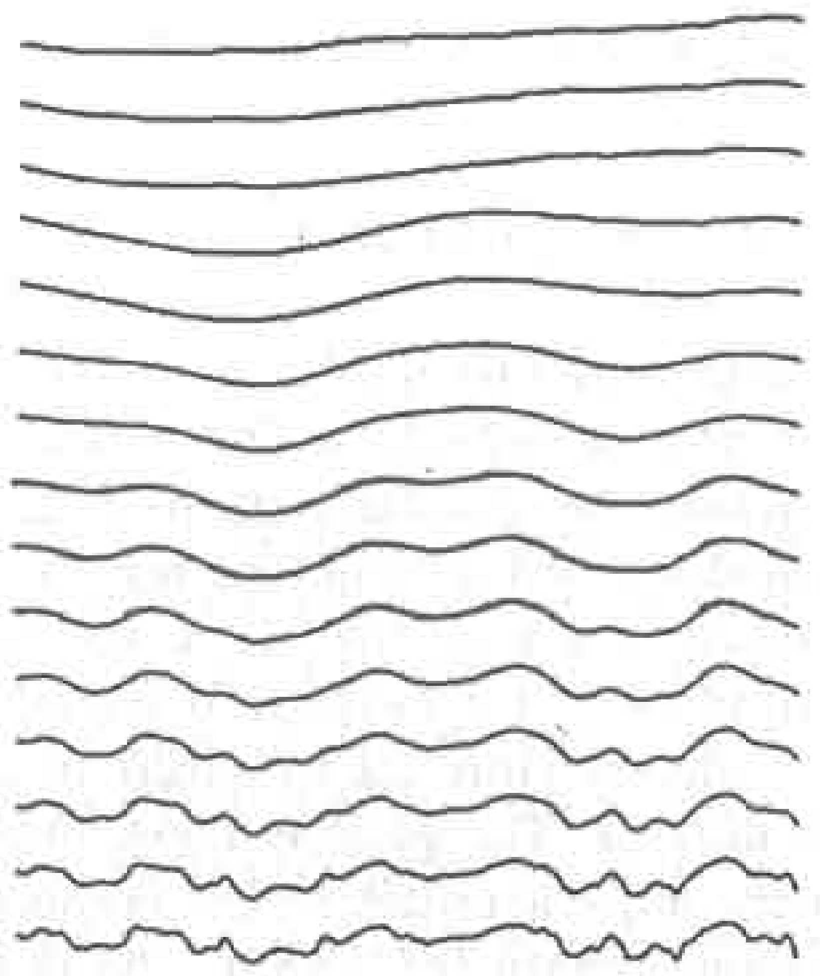

Figure 3 A sequence of gaussian smoothings of a waveform, with σ decreasing from top to bottom. Each graph is a constant-σ profile from the scale-space image.

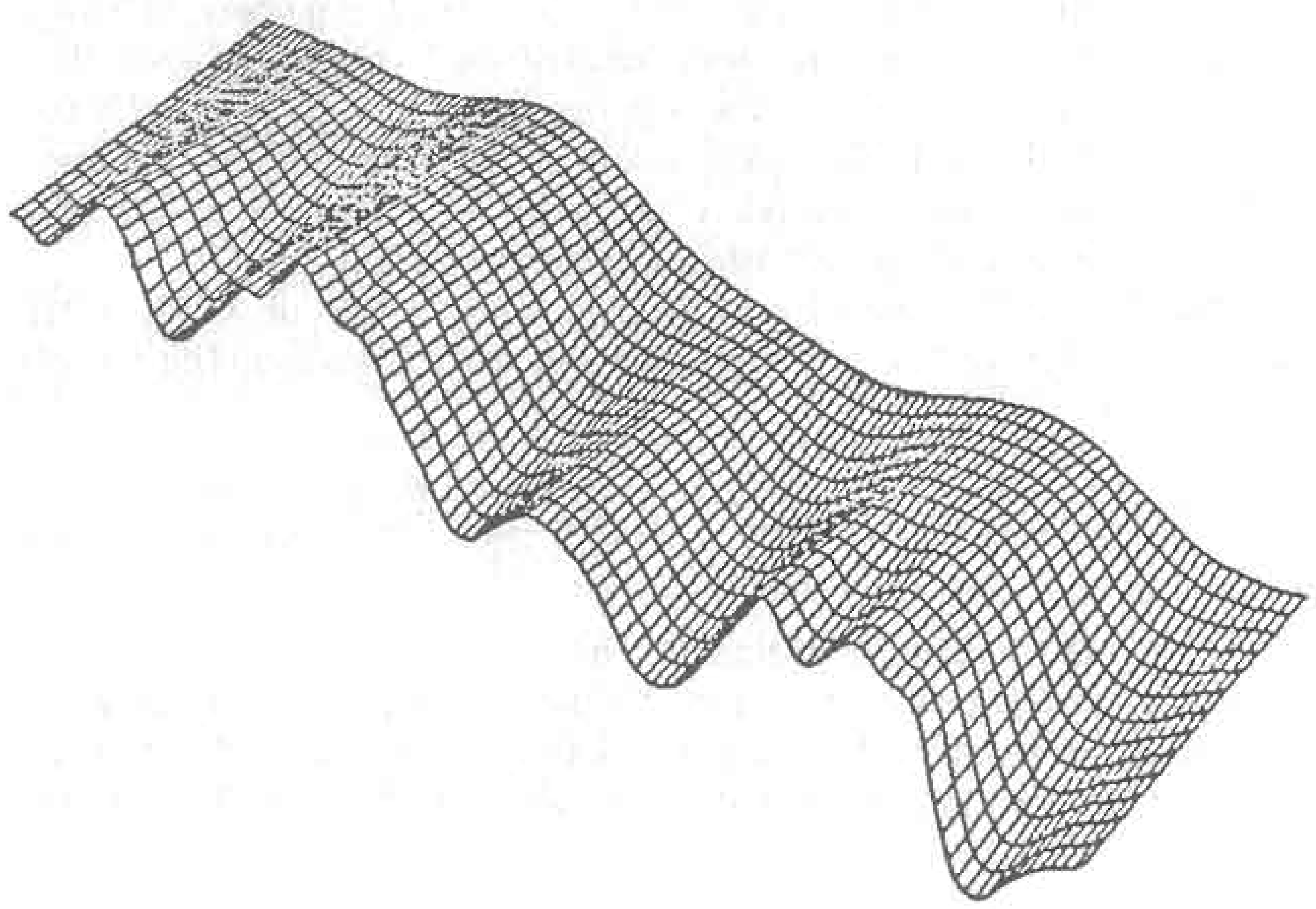

Figure 4 The same sequence of Figure 3 portrayed as a surface in perspective.

THE SCALE-SPACE IMAGE: QUALITATIVE STRUCTURE

An extremum in the nth derivative is a zero-crossing in the $(n + 1)$th. Although, conceptually, we are interested in extrema, working with the zero-crossings is more convenient. The nth order zero-crossings in F are the points that satisfy

$$\frac{\partial^n F}{\partial x^n} = f * \frac{\partial^n g}{\partial x^n} = 0, \quad \frac{\partial^{(n+1)} F}{\partial x^{(n+1)}} \neq 0 \tag{2}$$

where the derivatives of the gaussian are readily obtained.[2] These points are extrema in the $(n - 1)$th derivative. Thus, evaluating the partials at any fixed σ, the zero-crossings in $\partial F/\partial x$ are local minima and maxima in the smoothed signal at that σ, those in $\partial^2 F/\partial x^2$ are extrema of slope (inflections), and those in $\partial^3 F/\partial x^3$ are extrema of (unnormalized) curvature.

Sampling F along several lines of constant σ yields a filter-bank of the sort used by Marr and others. Were we to compute qualitative descriptions separately for each slice, we would face the basic scale problem discussed earlier: we would be confronted with a different description at each scale, having no clear basis for relating one to another, or deciding which to use when. This chaotic state of affairs is exactly what we wish to avoid.

The scale-space image, because it treats σ as a continuous variable, offers the means to do this. Figure 5 shows a typical waveform with the zero-crossings in $\partial^2 F/\partial x^2$ (inflection points) taken at several values of σ. Some of the inflections appear to correspond over scale, although they aren't exactly aligned.

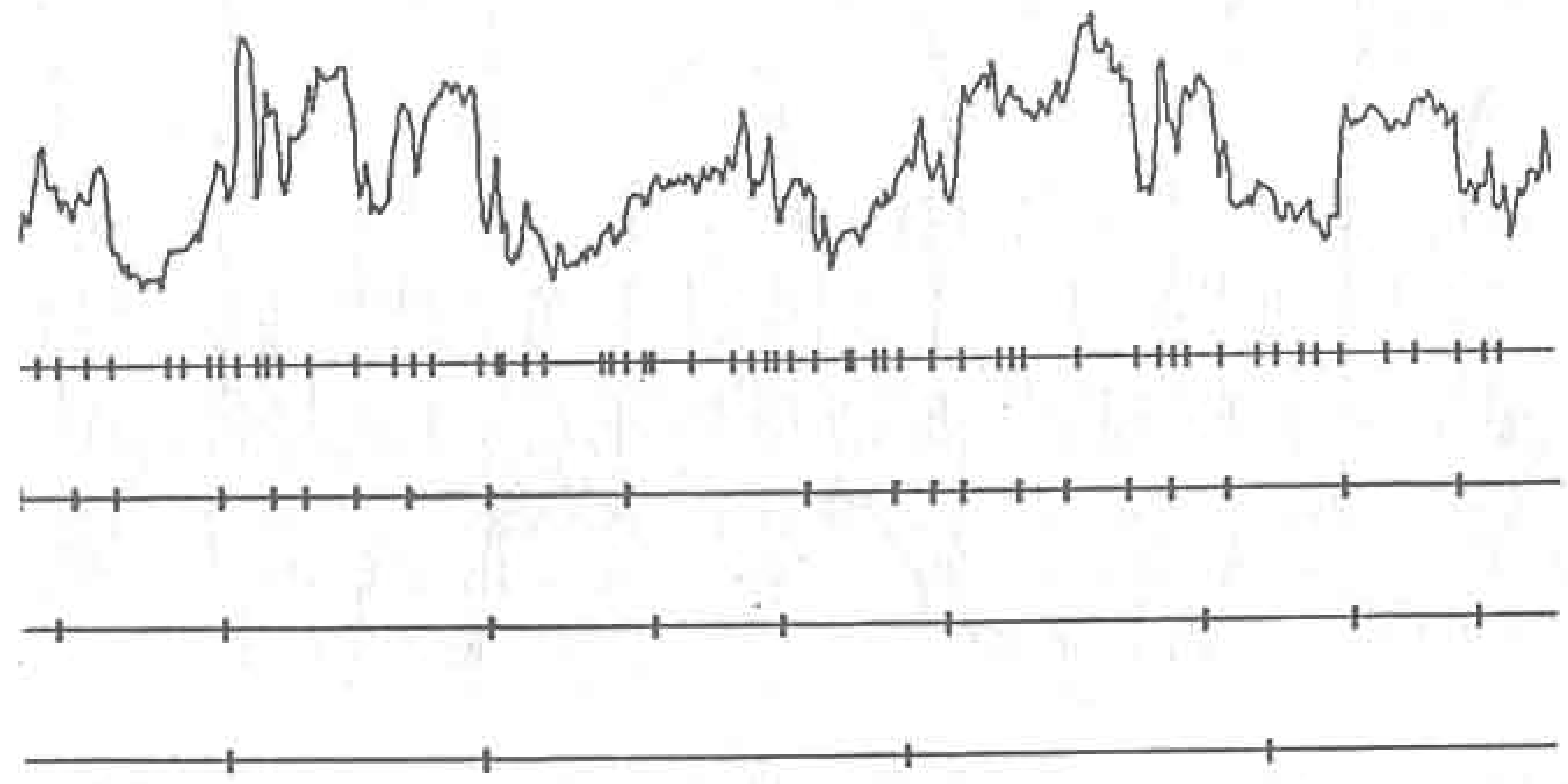

Figure 5 A typical waveform, with the locations of its second-order zero-crossings taken at several scales.

[2] We are interested only in the partials with respect to x, not σ, because only the former correspond to zero-crossings in the smoothed signal at some scale.

Others seem to have no correspondence, or ambiguous correspondence. Is there any meaningful sense in which certain extrema correspond over scale? Figure 6 shows what happens when we fill in the gaps: the fixed-scale zeroes in fact lie on zero-crossing *contours* through scale-space. Some of the isolated zeroes we observed in Figure 5 do correspond, in the sense that they lie on the same contour, while others do not.

Observe that these contours form arches, closed above but open below. At the apex of each arch, F satisfies

$$\frac{\partial^n F}{\partial x^n} = 0, \quad \frac{\partial^{n+1} F}{\partial x^{n+1}} = 0, \quad \frac{\partial^{n+1} F}{\partial x^n \partial \sigma} \neq 0, \qquad (3)$$

which is not a zero-crossing point by the definition in (2). Thus, by a "contour," we really mean a single arm of an arch, with the apex deleted. Each contour may be viewed as a zero-crossing that moves continuously on the x-axis as σ is varied, with instantaneous "velocity"

$$-\frac{\partial^{n+1} F/\partial x^n \partial \sigma}{\partial^{n+1} F/\partial x^{n+1}} \qquad (4)$$

The denominator in (4) is the slope at which $\partial^n F/\partial x^n$ crosses zero (the "strength" of the zero-crossing,) and the numerator responds to features entering or leaving the mask's effective receptive field. Thus, the location of strong zero-crossings

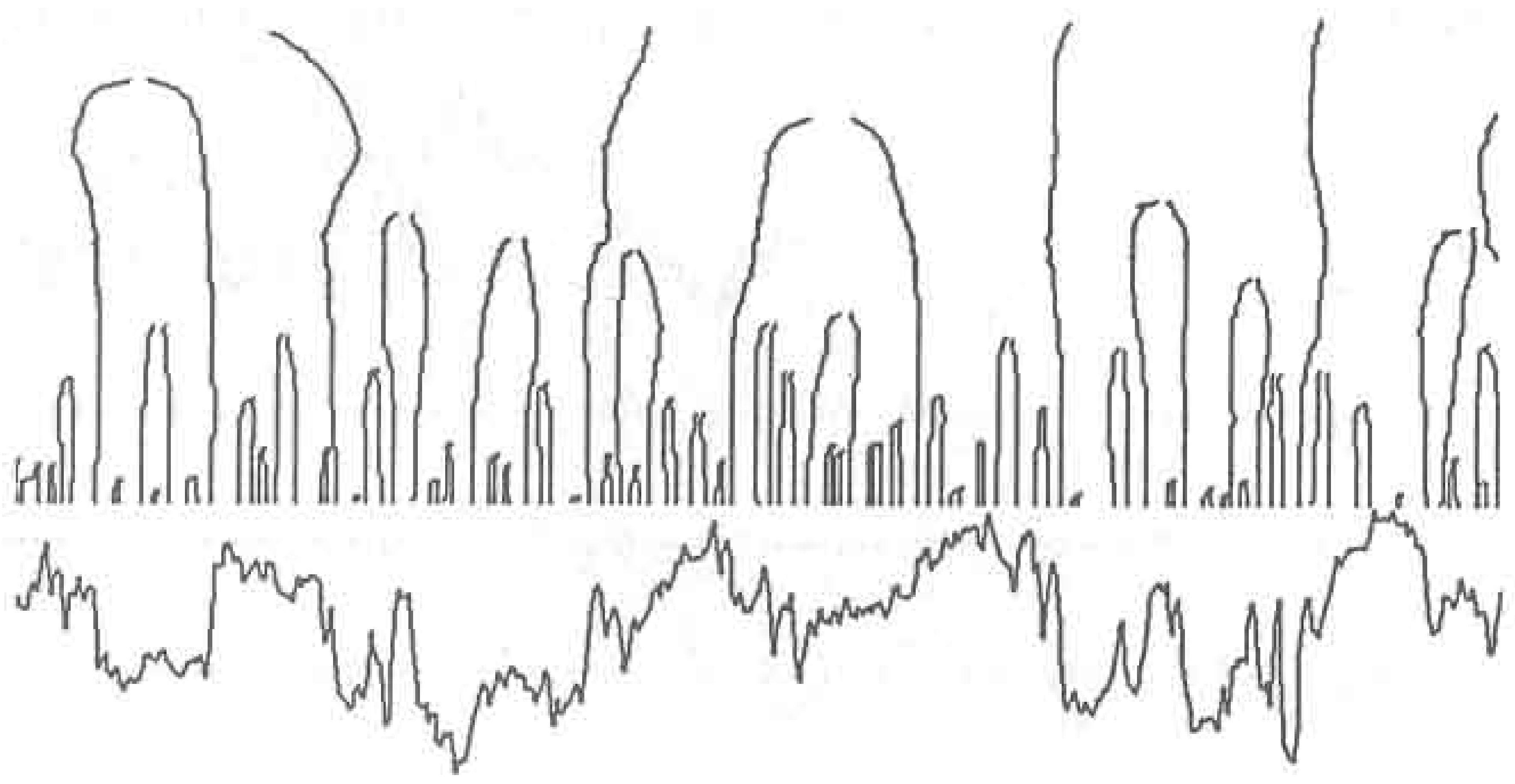

Figure 6 Contours of $F_{xx} = 0$ in the scale-space image. The x-axis is horizontal; the coarsest scale is on top. To simulate the effect of a continuous scale-change on the qualitative description, hole a straight-edge (or better still, a slit) horizontally. The intersections of the edge with the zero-contours are the zero-crossings at some single value of σ. Moving the edge up or down increases or decreases σ.

tends to be more stable over scale than that of weak ones, but any zero-crossing may be "pulled" by sufficiently large surrounding features. Also, as the strength of a zero-crossing approaches zero, its velocity may become arbitrarily large.

The two arms of each arch form a complementary pair, crossing zero with opposite sense. As we sweep across the apex of an arch, with σ increasing, the pair approach each other with increasing velocity, then collide and are annihilated. The proof of Babaud et al. (1983), mentioned above, assures us that the complementary singularity—a pair of zeroes vanishing as we move to a *finer* scale—can never occur.

THE "RAW" SCALE-SPACE DESCRIPTION

Thus, a decrement in σ has two effects on the zero-crossings: the continuous motion of existing zero-crossings, according to (4), and the appearance of new ones, in complementary pairs, at the singular points described in (3). A zero crossing observed at any scale may always be tracked continuously across all finer scales.

What do these properties imply for qualitative description? Suppose we qualitatively describe f at a particular scale by the sequence of zero-crossings at that scale, noting for each zero-crossing the sense with which it crosses zero, and the order of the extremum it represents (if more than one order is being used). Each of these fixed-scale zeroes lies on a zero-crossing contour in scale-space. The contours, once they appear, cannot disappear at finer scales, nor, in general, change their sense, nor cross each other. Therefore, the initial sequence must always be embedded, in its entirety, in the sequence obtained at any finer scale. The only changes in the qualitative description as σ decreases occur at the "apex" singularities, where a new pair of zero-crossings must be spliced into the sequence.

Therefore, we may capture the qualitative description at *all* scales just by recording the sequence of zero-crossing contours intersecting the finest observable scale, and for each, the value of σ at which the contour vanishes. The qualitative description at any scale is then obtained by deleting from the full sequence the zero-crossings that vanish at a finer scale. Rather than a collection of independent single-scale descriptions, this affords a unified one, each of whose elements exists over a definite range of scales. Complete quantitative information may be attached to the qualitative description by recording the contours' trajectories through scale space, and relevant properties of F along the trajectory.

Identity, Scale, and Localization Assumptions

We have defined the scale-space description, showing that it captures the qualitative structure of the signal at all scales. But what physical interpretation shall we place on the events comprising the description? What does a scale-space contour mean, what does the geometry of its trajectory tell us, and so forth?

We first assume that each contour reflects a single physical event. When we eventually say, for example, that "this inflection exists because . . ." we must complete the sentence in the same way for every point on the contour. This amounts to an assumption that the organization of the scale-space image is physically meaningful, that when we change σ a small amount we are almost always seeing the same broad-band events at a different scale rather than a whole new set of narrow-band events. This "identity" assumption allows us to treat each contour as an indivisible unit for the purpose of interpretation.

We next assume that the *scale* of each event, for the purpose of grouping and organization, is given by the "apex" singularity, the scale above which the contour vanishes. Thus, by a "large-scale event" we mean one that is observable at a large σ and all smaller ones. Finally, we assume that the true location of a zero-crossing contour on the x axis is its location at the finest observable scale. This assumption is motivated by the observation that linear smoothing has two effects: qualitative simplification–the removal of fine-scale features–and spatial distortion–dislocation, broadening, and flattening of the features that survive. The latter effect is undesirable, because a *big* event is not necessarily a *fuzzy* one. Because each event in the scale-space description extends over a range of scales, there is no difficulty in assigning a fine-scale location to a coarse-scale event. Collectively, these assumptions characterize each scale-space contour as denoting a single physical event, whose scale is the contour's vanishing scale, and whose location is the contour's fine-scale location.

Figure 7 illustrates the effectiveness of this interpretation: We select the events whose scale exceeds a threshold, then draw a "sketch" of the signal by fitting parabolic arcs between the distinguished points. This description is qualitatively isomorphic to the gaussian-smoothed signal at the same scale, but it much better preserves the locations of the distinguished points.

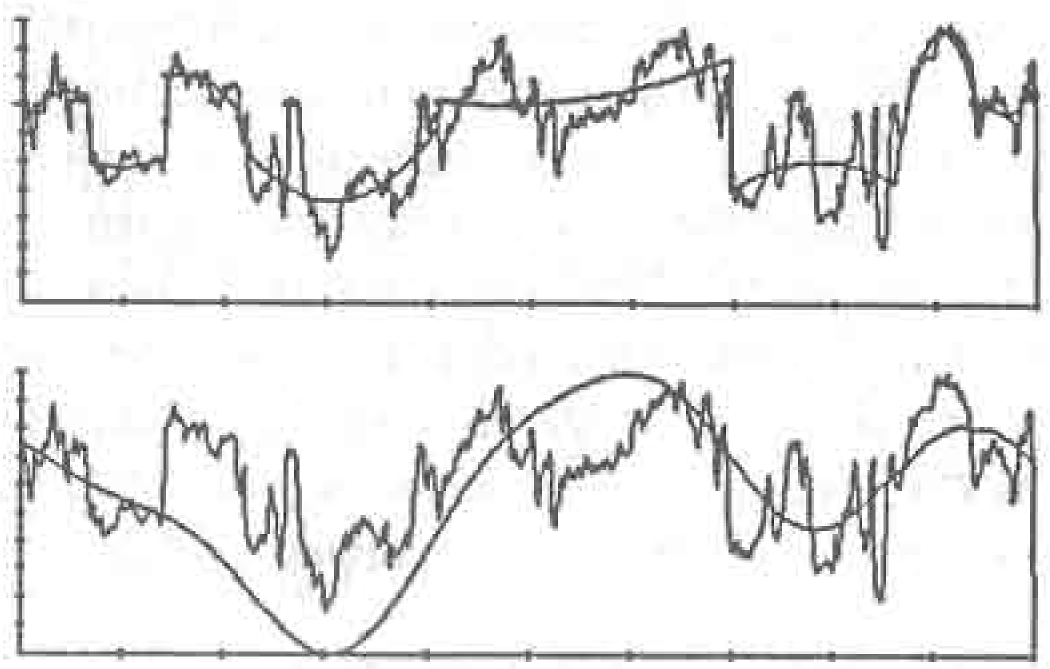

Figure 7 Above is shown a waveform with a superimposed approximation. The approximation was produced by identifying inflections at a coarse scale, and resolving their fine-scale locations. Parabolae were then fit independently between each pair of localized inflections. Below is shown the corresponding (qualitatively isomorphic) gaussian smoothing.

ORGANIZATION AND SCALE

We began by distinguishing the problem of *characterizing* primitive events over scale from the problem of grouping or *organizing* them in a meaningful way. The raw scale-space description, by design, is noncommittal with respect to organization: it concisely lists the primitive events observable at all scales, each with its scale and location, as well as other qualitative and quantitative information. Given the raw description, one is free to use the events' scale, as well as the rest, in whatever manner one likes. In this section, we will introduce a representation, derived from the raw scale-space description, that is organized by scale. This representation, called the interval tree, describes the division of intervals bounded by extrema into finer and finer sub-intervals, as σ decreases. The scale channels proposed by Marr and others produce poor organizations because events are structured by a series of arbitrary global scale thresholds. In contrast, the interval tree permits the scale of description to be controlled in local discrete steps, determined by the qualitative structure of the signal.

The Interval Tree

The scale-space extrema whose scales exceed a given threshold, σ_T, partition the x axis into intervals. As we decrease σ_T, starting from a coarse scale, new events appear in pairs (each associated with an "apex" singularity in the scale-space image), dividing the enclosing interval into a triple of sub-intervals (see Figure 8). As σ_T decreases further, these new intervals in turn subdivide, and so on down to the finest observable scale.

Each of these intervals defines a rectangle in scale-space, bounded above by the scale at which the interval emerges out of an enclosing one, bounded below by the scale at which it splits into sub intervals, and bounded on either side by the x-locations of the events that define it. Thus each interval has a definite

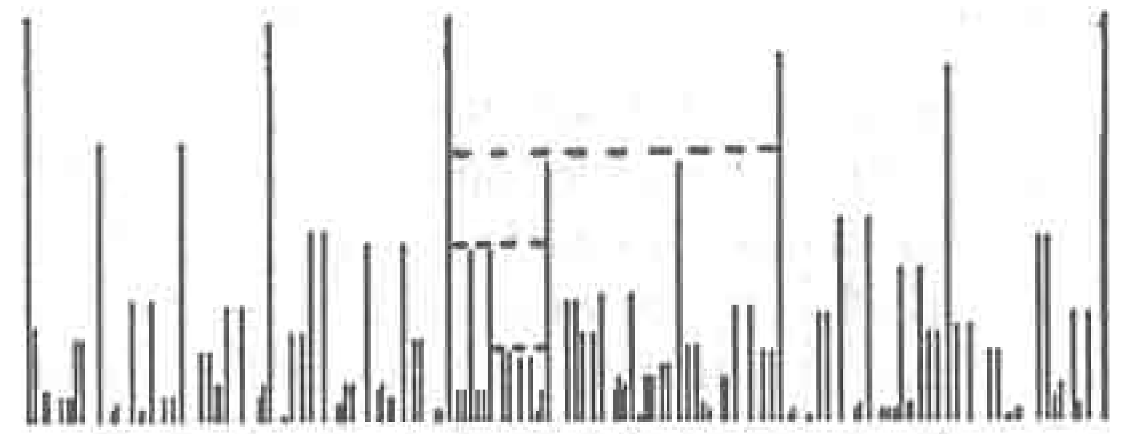

Figure 8 The vertical lines represent a sequence of distinguished points, with the height of each line representing the scale at which the event vanishes. The events whose scale exceeds a threshold, σ_T, partition the x-axis into intervals. As we decrease σ_T, new events appear in pairs, dividing the enclosing interval into a triple of sub-intervals.

width and location on the signal axis, and exists over a definite range of scales. Collectively, the rectangles tesselate the (x, σ) plane.

The intervals also correspond to nodes in a ternary-branching tree: an interval's parent is the larger interval from which it emerged, and its offspring are the sub-intervals into which it divides. This *interval tree* is illustrated in Figure 9. (For convenience, we take as the root of the tree a dummy interval including the whole signal).

Using the Interval Tree

The coverings of the x-axis by the interval tree[3] each segment the signal into a set of intervals. From any starting point, we may generate a new segmentation either by splitting an interval into its offspring, or merging some intervals into their common parent. We may thus explore the space of descriptions, changing the scale of description in local, discrete steps (see Figure 10). We have, in effect, collapsed the (x, σ) plane into a discrete set, taking advantage of the singular points at which extrema appear to do so.

The set of segmentations generated by the interval tree is a small fraction of the segmentations one could generate by collecting arbitrary subsets of all the scale-space extrema. In particular, a segmentation is excluded if any of its intervals contains an extremum of larger scale than either of the two bounding ones. This "interval constraint" is in fact a useful and conservative one, because it discards the myriad of meaningless groupings—constructed, for example, by joining pairs of widely separated fine-scale extrema, or skipping an arbitrary number of large-scale ones along the way—while preserving "smooth" intervals that are broken only by extrema of much finer scales than the bounding ones. We found, by extensive but informal testing, that people are nearly always able to duplicate the segmentations they find perceptually salient within this constraint, by moving interactively through the tree. Figure 11 shows a few signals with "sketches" that have been obtained in this way.

The Stability Criterion

Although the interval tree generates descriptions in an organized way, it would be useful to establish a preference ordering on those descriptions, to obtain "top level" sketches as starting points for matching and interpretation. The interval tree initially requires that all extrema within each interval be of finer scale than the bounding ones. A natural extension to this useful constraint would be to favor intervals for which the difference between the scale of the bounding extrema, and that of the largest extremum between them, is large. In the interval tree, this difference is just a node's extent in the scale domain—its persistence or

[3] That is, all the sets of nodes for which every point on the x-axis is covered by exactly one node.

stability. A very stable node is one bounded by very large-scale extrema, but containing only very small ones—intuitively, a "smooth" interval.

We obtain *maximum stability descriptions* by starting from the root of the tree, and moving down recursively whenever a node's stability is less than the mean stability of its offspring. These are local maxima, which may be layered at successively finer scales. (Several variants on this procedure achieved similar results.) Figure 12 shows several signals with their top-level maximum stability descriptions. The reader should compare these with their own percepts, and also observe that they closely resemble the interactively obtained descriptions of Figure 11.

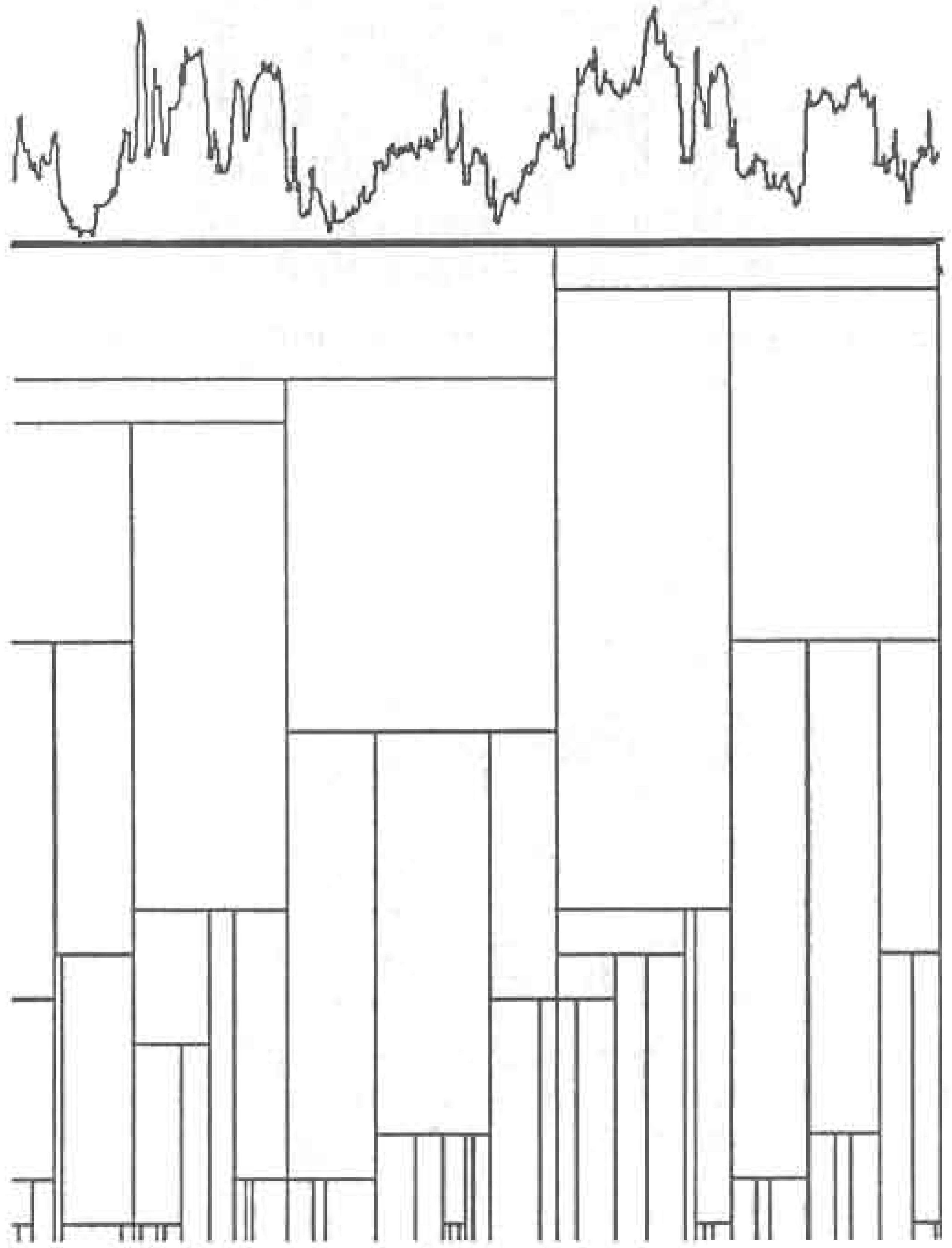

Figure 9 A signal with its interval tree, represented as a rectangular tesselation of scale-space. Each rectangle is a node, indicating an interval on the signal, and the scale range over which the signal interval exists.

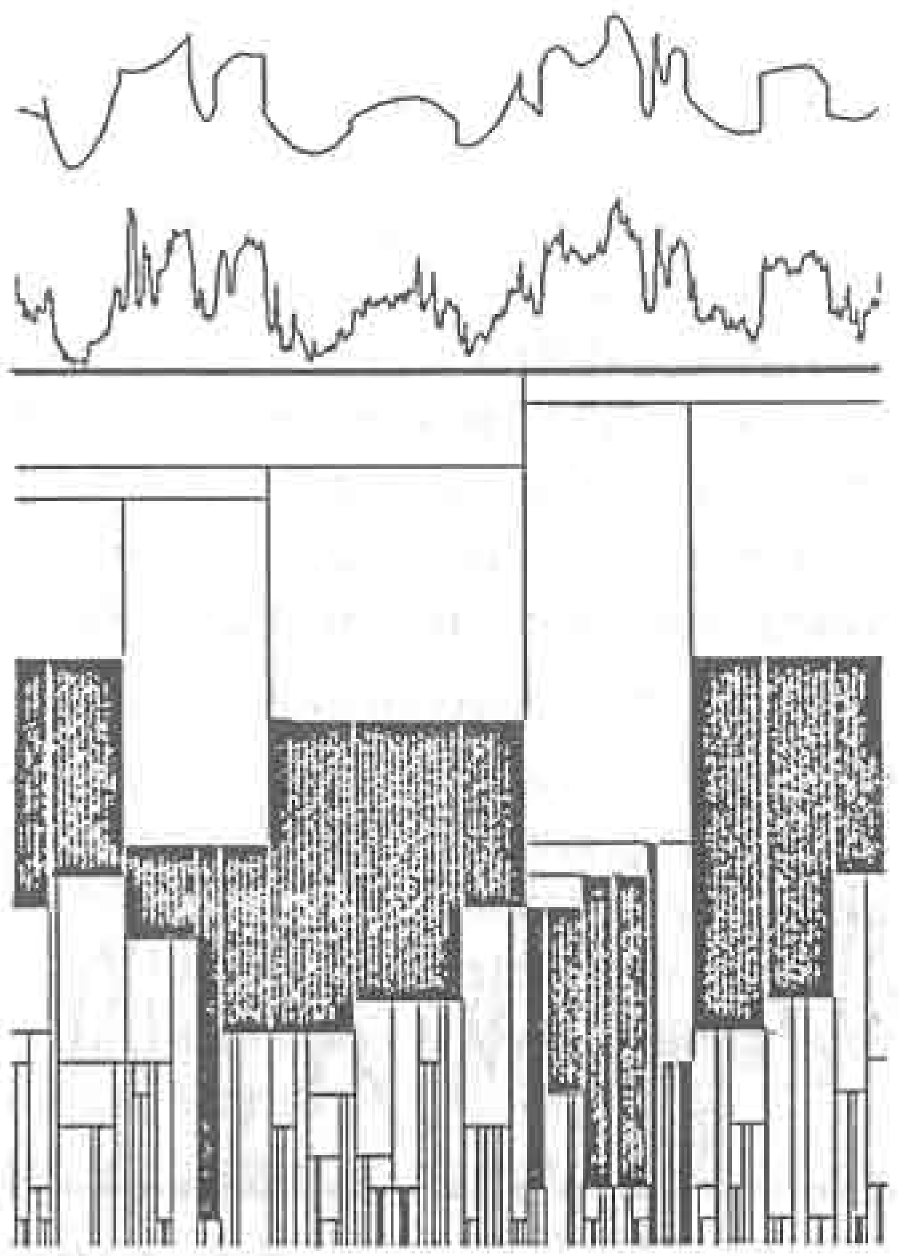

Figure 10 An approximation derived from a particular covering of the signal by the interval tree. The "active" nodes (i.e., those that determined the segmentation) appear in black.

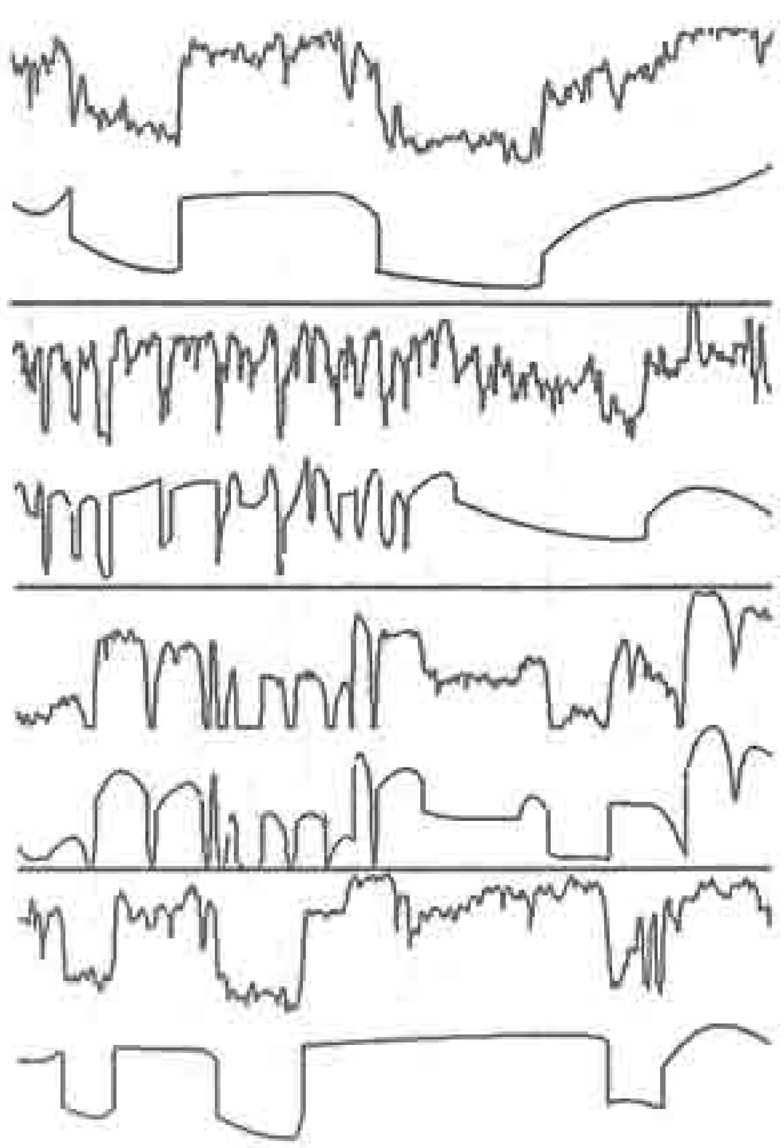

Figure 11 Descriptions obtained by interactively traversing the interval tree to match perceptually salient segmentations.

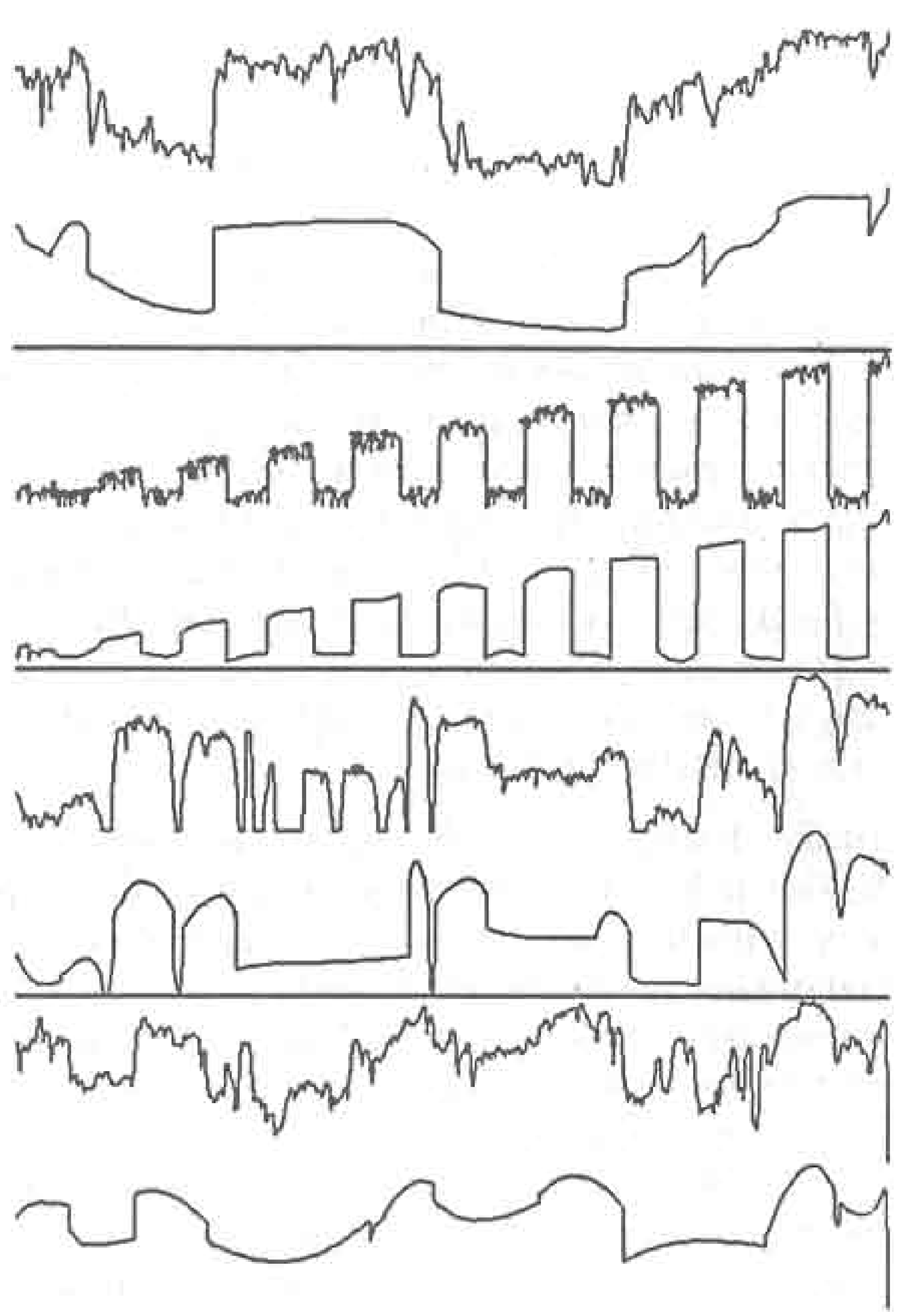

Figure 12 Several signals, with their maximum-stability descriptions. These are "top-level" descriptions, generated automatically, and without thresholds. The reader should compare the descriptions to their own first-glance "top-level" percepts.

CONCLUSIONS

We set out with two objectives: first, to obtain a unified description of the extrema in a signal over a wide range of scales; and second, to use the scale of extrema to constraint and guide their organization and grouping. Both objectives were attained.

Description: The extrema observed at a particular scale move continuously as the scale is varied, vanishing in pairs at singular points as the scale becomes coarser. These continuously moving extrema form contours on the scale-space image, the surface swept out by smoothly varying the scale. We take these contours as the units of description, on the assumption that each in general denotes a single physical event. The qualitative structure over all observed scales is captured in a single sequence of extrema, each extending over a range of scales, and each characterized by its fine-scale location and the scale above which it vanishes.

Organization: The least unit of organization is an interval bounded by a pair of extrema. A criterion for meaningful groupings is the difference in scale between the bounding extrema and the ones within the interval: where the bounding extrema are much larger, the interval is a strong candidate for treatment as an "undistinguished interval" (i.e., monotonic in the derivatives in which extrema were found). This grouping criterion is captured in the interval tree, a ternary-branching structure, simply derived from the bare description that describes the coarse-to-fine subdivision of intervals. The structure of the tree itself excludes intervals containing larger-scale extrema than the bounding ones. Within this constraint, the persistence or stability over scale of an interval in the tree measures the scale difference. Using this measure, maximum stability groupings were constructed to provide "top-level sketches" of the signal.

The bare scale-space description, and the organization provided by the interval tree, constitute a sound basis for further organization, by discovering regularities such as symmetry, repetition, and inter-signal correspondence; for the construction of larger morphological structures expressed as sequences of extrema; and for graphic communication between man and machine. We are currently investigating all of these topics, as well as applying the basic methods to signal understanding problems in geology, vision, and speech.

An extension of these methods to two-dimensional images is also under way. The two-dimensional world is richer and more complex—the primitive events are not points but contours, and their extensions in scale-space are surfaces in a volume. Nevertheless, the elements of the one-dimensional method carry over quite directly.

REFERENCES

Babaud, J., Witkin, A., & Duda, R. (1983). *Uniqueness of the gaussian kernel for scale-space filtering.* (Fairchild TR 645).

Ballard, D. H., & Brown, C. M. (1982). *Computer vision.* Englewood Cliffs, NJ: Prentice Hall.

Haralick, R. M., Watson, L. T., & Laffey, T. J. (1982). *The topographic primal sketch* (Technical Report CS Department). Blacksburg, VA: Virginia Polytechnic Institute.

Hoffman, D. (forthcoming). *Representing shapes for visual recognition.* Unpublished doctoral dissertation, MIT, Cambridge, MA.

Hong, T. H., Shneier, M., & Rosenfeld, A. (1981). *Border extraction using linked edge pyramids* (TR-1080, Computer Vision Laboratory). College Park, ML: University of Maryland.

Mandelbrot, B. (1977). *Fractals: Form, chance, and dimension.* San Francisco: W. H. Freeman.

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Marr, D., & Hildreth, E. C. (1979). *Theory of edge detection* (M.I.T. Artificial Intelligence Memo 518). Cambridge, MA: MIT.

Marr, D, & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London*, B204, pp. 301–328.

Pavlidis, T. (1977). *Structural pattern recognition.* New York: Springer.

Pentland, A. P. (1983). Fractal-based description. *Proceedings of the ICJAI.*

Rosenfeld, A., & Kak, A. C. (1976). *Digital picture processing.* New York: Academic Press.

Rosenfeld, A., & Thurston, M. (1971). Edge and curve detection for visual scene analysis. *IEEE Transactions on Computers*, C-20, pp. 562–569.

Witkin, A. P. (1983). What is perceptual organization for? *Proceedings of the IJCAI.*

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In A. Rosenfeld (Ed.), *Human and Machine Vision.* New York: Academic Press.