

Reconstructions from Zero Crossings in Scale Space

ROBERT HUMMEL, MEMBER, IEEE, AND ROBERT MONIOT

Abstract—A useful representation of signal data, besides being a complete and stable transformation of the information, should make explicit important features in the data. In computer vision, the one-parameter family of images obtained from the Laplacian-of-a-Gaussian-filtered version of the image, parameterized by the width of the Gaussian, has proven to be a useful data structure for the extraction of feature data. In particular, the zero crossings of these so-called scale-space data are associated with edges, and were proposed by Marr and others as the basis of a representation of the image data. The question arises as to whether the representation is complete and stable. We survey some of the results and studies related to these questions, and survey several papers that attempt reconstructions based on this or related representations. We then formulate a new method for the reconstruction from zero crossings in scale space, based on minimizing equation error, and present results showing that the reconstruction is possible, but can be unstable. We further show that the method applies when gradient data along the zero crossings are included in the representation, and demonstrate empirically that the reconstruction is then stable.

I. NOTIONS

A. Representations

IN the fields of signal analysis and image processing, the first stage of a complete pattern recognition system typically applies some numerical process to the digital data. In image processing, for example, it is generally considered useful to extract edges, corners, and textured regions in the image. In other words, the features and salient symbolic information about signal and image data are dynamically associated with groups or regions of the data, and typically depend more explicitly on primitive features such as discontinuities, extrema, and local statistical properties of the spatially sampled data. It seems reasonable, therefore, to transform the initial data to intermediate representations to make the primitive features more accessible to the algorithms for signal analysis. The collection of all intermediate representations, which might include binary edge images, texture measures, and other feature detectors, comprise a representation that replaces the original signal for analysis purposes. This is the central idea in vision processing, for example, of Marr's "Primal Sketch" [1] or Tenenbaum and Barrow's "Intrinsic Images" [2].

Since all analysis is done on the intermediate represen-

tation, the representation should carry all the information necessary for the interpretation of the data. Of course, the main idea is that the relevant information should be more explicit than the original samples, and that data redundancies should be eliminated. The representation may constitute a data compressed version of the original sampling, but might as easily contain more bulk data, in the attempt to represent different features. It should be possible to reconstruct a version of the original signal from the intermediate representation since the representation should contain all the essential information. While the reconstruction need not be mathematically identical, it should have the same subjective interpretation as the original signal.

It is important to emphasize that our interest, from the computer vision standpoint, is independent of data compression concerns. An intermediate representation such as the zero crossings together with the gradients along the zero crossings, discussed in detail later in this paper, may contain redundant information and fail to be compressive at all. Instead, our concern is with the quality of an intermediate representation for the purposes of building a computer vision system. Such a system will probably never actually perform a reconstruction. Accordingly, we do not care if the representation is bulky and if reconstruction is expensive. We *do* care that the construction of the representation should be easy, and that stable reconstruction is at least theoretically possible.

How can we evaluate the quality of an intermediate representation? The usual method is to show that algorithms making use of the representation are effective. While it is difficult to fault success, this approach is limited by a lack of generality. That is, the range of algorithms that can be demonstrated is necessarily limited, and the reasons for their effectiveness and limitations are always varied. For example, we can never be sure if a partly successful vision system is limited because the features are inadequate or because the pattern matching method is inextensible.

Alternatively, we can study a representation mathematically, to understand 1) the dependence of the representation on the initial data, 2) the fibers of the representation, and 3) the stability of the transformation. For the first issue, the dependence, it is desirable to show some form of continuity, so that small changes in the original data result in small changes in the representation. By "fibers," as studied in the second issue, we mean the sets of signals that map to a single representation. If the mapping is one-to-one, then the fibers are singleton sets, and the representation is complete. Finally, stability of the

Manuscript received March 17, 1989. This work was supported by NSF Grant IRI-8703335 and Office of Naval Research Grant N00014-85-K-0077, Work Unit NR 4007006.

R. Hummel is with the Courant Institute of Mathematical Science, and the Washington Square Center for Neural Science, New York University, New York, NY 10012.

R. Moniot is with the Division of Science and Mathematics, College at Lincoln Center, Fordham University, New York, NY 10023.

IEEE Log Number 8931326.

applications. Many of the operators assumed a certain resolution level, and apply the operations to a neighborhood of samples that is independent of the data. Thus, many edge detectors assume that the edge can be discerned within a three-by-three pixel neighborhood. Such operators frequently fail to find features when the data are sampled too finely, and sometimes fail when the resolution is too coarse. Nonetheless, in controlled cases, feature operators can perform useful preprocessing for the analysis of signal data.

In order to build into the process a certain degree of scale invariance, and to make the feature detection operate dynamically on the data, the use of multiresolution data structures has become an important aspect of signal processing. One of the earliest uses of multiresolution methods in computer vision was described by Rosenfeld and Thurston [4]. A simple construction is to sample the image at multiple resolution levels, forming a "pyramid" of images. Each layer represents a different scale for the same image. Algorithms are then developed that incorporate feature-detector responses obtained by applying the same operation to each level of the pyramid. From a single high-resolution image, this pyramid can be obtained numerically by successively blurring and subsampling the original image. Each operation of blurring and subsampling produces an antialiased version of the same image at a coarser resolution. The resulting structure is called a "Gaussian pyramid." Computational methods for building Gaussian pyramids, and other pyramids, are described by Burt and Adelson [5].

We will describe a continuous formulation of these multiresolution pyramids, in order to assist in the mathematical analysis, and to formulate scale space precisely. We begin with data $f(x)$, where $x \in \mathbb{R}^n$. For $n = 1$, $f(x)$ could be an acoustical signal; for $n = 2$, we usually regard $f(x)$ as an image. The case $n = 3$ arises with seismographic data, time-varying imagery, or tomographic data. Higher dimensional domains for the data are possible. *Scale space* refers to a domain (x, t) for a set of data, $u(x, t)$, parameterized by a variable $t \in \mathbb{R}$, giving variations on the data $f(x)$. In particular, we will always assume that $t \geq 0$, and that $u(x, 0) = f(x)$. Further, $u(x, t)$ will be continuous in t , and that for any $t_0 > 0$, $u(x, t_0)$ gives a variant of $f(x)$. The idea of scale space, parameterized by a continuous variable obtained as the standard deviation of a Gaussian, is due to Witkin [6].

The natural framework for the analysis of scale-space formulations of multiresolution representations is in terms of the heat equation [7], [8]. We define $u(x, t)$ to be a bounded solution to the heat equation:

$$\frac{\partial u}{\partial t} = \Delta u,$$

$$u(x, 0) = f(x). \quad (\text{Heat Equation})$$

The solution is given by convolution against the fundamental solution to the Heat Equation, which for the do-

main \mathbb{R}^n is given by

$$u(x, t) = \int_{\mathbb{R}^n} K(x - y, t) f(y) dy,$$

where

$$K(x, t) = (4\pi t)^{-n/2} e^{-|x|^2/4t}.$$

We see that $u(x, t)$ is obtained by blurring $f(x)$ by increasingly diffuse Gaussians, parameterized by $t > 0$, with standard deviations σ satisfying $2\sigma^2 = 4t$. In computer vision, scale space sometimes refers to the (x, σ) variables that can be used to reparameterize the domain of u . We retain the (x, t) parameterization to keep the linear Heat Equation relation for the function u .

Convolution by Gaussians is considered special for many reasons [7], [9], [10]. We see from the above analysis a relationship between Gaussian convolution, the Heat Equation, and the Laplacian operator. Of course, Gaussian convolution enjoys other properties; for example, the central limit theorem implies that Gaussian convolution is easy to implement by an iterative procedure.

In computer vision, in an idea that dates back to 1955 [11], the image data $f(x)$ are often filtered with the Laplacian of a Gaussian, instead of filtering by a Gaussian. The use of a Laplacian-of-Gaussian-type filter is in part motivated by the existence of center-surround receptive fields, which exhibit excitatory/inhibitory center/surround response patterns [12], but can also be motivated on a predictive encoding basis [13]. Experiences with Laplacian-of-Gaussian convolutions have demonstrated their utility for computer vision applications. Filtering by the Laplacian of a Gaussian can be written in three ways:

$$\Delta K * f = K * \Delta f = \Delta(K * f).$$

If we denote the result by $v(x, t)$, we see that

- 1) $v(x, t)$ is the $f(x)$ data filtered by the Laplacian of a Gaussian,
- 2) $v(x, t)$ is the solution to the Heat Equation with initial data Δf , and
- 3) $v(x, t)$ is $\Delta u(x, t)$, where u is the solution to the Heat Equation with initial data $f(x)$.

We will refer to v as the scale-space function of the data f . The scale-space function v provides a continuous-domain analog [14] of the Laplacian-pyramid data structure [5].

A difference of two Gaussians is often used in place of the Laplacian-of-Gaussian. We see from the Heat Equation formulation the basis for the approximation. Namely, since $K(x, t)$ is itself a solution to the Heat Equation,

$$(\Delta K)(x, t) = \frac{\partial K}{\partial t}(x, t)$$

$$= \lim_{\tau \rightarrow 0} (K(x, t + \tau) - K(x, t))/\tau.$$

That is, the difference of Gaussians is a good approximation to ΔK as the separation between the spread of the

two Gaussians approaches zero (and the difference is scaled). In actual use, it is more frequent to use a difference of Gaussians where the ratio of the two scales in the σ -variable is given by 1.6, but the rationale for this choice is not the degree of approximation to the Laplacian-of-Gaussian [15].

C. Zero Crossings

The zero set of $v(x, t)$ is the point set in (x, t) where $v = 0$. The set might be empty (for instance, if f is subharmonic or superharmonic [16]) everything (if f is harmonic [17]), or a proper subset of (x, t) space. In the latter case, zeros can be isolated points, lines, and surfaces (but never regions). We distinguish components of the zero set which form manifolds of codimension one.

Definition: The zero crossings of $v(x, t)$ refer to the point set

$$\partial\{(x, t) | v(x, t) < 0\} \cap \partial\{(x, t) | v(x, t) > 0\},$$

where ∂S refers to the boundary of the set S . ■

That is, the zero crossings are surfaces in scale space that separate regions where the scale-space function v is positive from regions where v is negative. If we intersect these zero-crossing surfaces with a plane $\{t = t_0\}$, fixing a single level, we obtain the "zero-crossing curves" at that level [see Fig. 1(b)]. These curves also separate positive and negative regions, but in this case the regions lie in the plane determined by the fixed level. The zero-crossing curves at the continuum of levels trace, and thus develop, the zero-crossing surfaces, which we refer to simply as the zero crossings. Of course, our terminology of "curves," "surfaces," and the fixed "plane" are motivated by the case of a two-dimensional spatial domain (i.e., $x \in \mathbb{R}^2$); however, the concepts are valid for any spatial dimension.

Zero-crossings curves have been used for segmentation of imagery by edge detection [15], and for stereo matching and motion correspondence between pairs of images (e.g., [18]). It has also been suggested [10] that they form a nearly complete representation of Δf . Marr also suggested that the zero crossings might be a complete representation [1], but further thought that perhaps the zero-crossing information would need to be supplemented with gradient data along the zero crossings. These suggestions form the basis of the interest in zero crossings as an intermediate representation for images.

It should be noted that there is a long and distinguished history of the analysis of zero crossings in mathematics. Kedem supplies a survey together with some advances [19]. In this and related work, analysis is generally restricted to signals of one variable, and the representation is based on the local statistics of zero crossings (the number of crossings per unit length). Issues such as the local density of zero crossings become very important when they are used for stereopsis computation with image pairs. Presumably, this body of mathematical work is quite relevant to image representation, but has not been applied to computer vision in a major way.

Witkin [6] observes that zero crossings in scale space evolve as t increases, and are never created at some non-zero t . This property, discussed in [9] and in [10], ensures that zero-crossing surfaces are nested, one within another, enclosing regions containing the face $\{t = 0\}$, or forming a sheet meeting the face $\{t = 0\}$ and extending to $t = \infty$. In Fig. 1(b), this property is reflected in the fact that the zero-crossing surfaces close at the top—there are no surfaces that hang from the top, closing below. The property can be given a precise statement.

Evolution Property of Zero Crossings: Let C be a connected component of the set of zero crossings in the domain $\{(x, t) | x \in \mathbb{R}^n, T_1 \leq t \leq T_2\}$, where $0 \leq T_1 < T_2$. Then $C \cap \{(x, t) | t = T_1\} \neq \emptyset$. ■

When first observed, the evolution property sparked a great deal of interest in scale-space methods. It was thought that the property gave great justification to the use of Laplacian-of-Gaussian filtering, and indeed, that the Gaussian was the only filter that would yield this property [9], [10]. Under appropriate restrictions, it is true that the Gaussian is special in terms of yielding the evolution property. However, as we will see in Section II-A, the evolution property is essentially an expression of the Maximum principle for the Heat Equation.

The practical import of the evolution property is that the zero-crossing information can be simplified, and represented in a symbolic (or approximate) way by describing the nesting and some other simple features of the zero-crossing surfaces. The hope is that the simplified or symbolic representation will also suffice to form a complete representation of the original data, at least for image analysis purposes. For instance, Johansen considers the representation of signals by the "toppoints" of the zero crossings in scale space (with t extended to $-\infty$) [20]. To date, the hope of simplifying the representation of the fingerprint of zero crossings has borne little practical fruit, but has nonetheless sparked a wide range of interesting mathematical analysis.

In the next section, we consider some of the mathematical results related to the evolution property and the issue of completeness of the representation by zero crossings.

II. ANALYTICAL RESULTS

A. Evolution Property

In this section, we show that the evolution property, applied to level crossings (and not just zero crossings), is equivalent to the classical Maximum Principle for parabolic partial differential equations. That is, we will show that the evolution property can be derived simply from the maximum principle, and moreover, that any scale-space construction that obeys the evolution property for all level crossings, subject to minor restrictions, will satisfy a maximum principle.

The classical maximum principle for the solution to the parabolic Heat Equation $\partial u / \partial t = \Delta u$ states (see, e.g., [21]–[23]):

Maximum Principle: Let $D \subseteq \mathbb{R}^n$ be open and bounded. Suppose u is a solution to the Heat Equation in $T = \{(x,$

$t) | x \in D, 0 < t < T \}$ of class C^2 and is continuous in the closure \bar{T} . Then u assumes its maximum at some point (x, t) for which either $x \in \partial D$ or $t = 0$.

The maximum principle holds for more general parabolic partial differential operators, but certainly not for all parabolic equations. It is not known (to the authors' knowledge) how to characterize parabolic equations giving a maximum principle. For example, the maximum principle does not hold when the biharmonic operator is used: $\partial u / \partial t = \Delta^2 u$. Generally, the elliptic operator will be second order in order for a maximum principle to hold.

We will assume that whenever a maximum principle holds, a similar minimum principle also is given. For linear parabolic operators, such as the Heat Equation, a minimum principle follows from the maximum principle. However, since we can envision more general operators, we will henceforth assume that the statement that a maximum principle holds means that both a maximum and minimum principle hold.

Next, let us denote the construction of the representation of the image f in scale space by the operator $v = Sf$. In the previous section, we described the construction of v , given $f \in L^\infty(\mathbb{R}^2)$, according to

$$v(x, t) = \int_{x,2} \Delta K(y - x, t) f(y) dy.$$

However, we now wish to permit more general constructions, allowing, for example, f to be defined on an irregular or bounded domain, or allowing constructions involving differential operators with nonconstant coefficients. However, we will need to place restrictions on the growth of functions $v = Sf$ as $|x| \rightarrow \infty$ and on the boundary of the domains. Specifically, we will assume that the domain of f is D (which may or may not be bounded), and that the corresponding scale space will have $(x, t) \in D \times [0, \infty)$. There is a class of permissible images f for which the scale-space operator $v = Sf$ is defined, and we assume that for v 's constructed in this fashion:

- i) $v \in C(D \times (0, \infty))$
- ii) $v(x, t) = 0$ for $x \in \partial D, t \geq 0$;
- iii) $v(x, t) \rightarrow 0$ as $|x| \rightarrow \infty$ uniformly for $t \geq 0$.

Note that this means for the Laplacian-of-Gaussian construction given above, we must limit application to functions f that tend to zero as $|x| \rightarrow \infty$. The space $f \in L^2(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$ suffices, for example. This restriction is stronger than necessary, but not terribly objectionable.

Our main result in this section is that a scale-space construction $v = Sf$, given the boundary restrictions, satisfies the evolution property for level crossings if and only if a maximum principle holds (meaning both a maximum and minimum principle). The result is Theorem 1.

Theorem 1: Consider a scale-space construction $v = Sf$ satisfying i), ii), and iii) above for all admissible functions f . Then the following are equivalent.

- 1) The maximum principle holds for functions $v = Sf$ obtained from admissible f 's.
- 2) The evolution property holds for level crossings of solutions $v = Sf$ obtained from admissible f 's.

Proof: We first show that the maximum principle implies the evolution property. For if the evolution property fails for a level crossing l of $v(x, t), v = Sf$, then there is a component C , and bounds t_1, t_2 , where C is a connected component of $\{v(x, t) > l | t_1 \leq t \leq t_2\}$ or $\{v(x, t) < l | t_1 \leq t \leq t_2\}$ such that C does not meet the plane $\{t = t_1\}$. If $l \neq 0$, then the component C must be bounded, since $|v| \rightarrow 0$ uniformly as $|x| \rightarrow \infty$. Thus, the supremum (or infimum) of $v(x, t)$ is attained at a finite point (x_0, t_0) in C . If $l = 0$, then the supremum (or infimum) in C is nonzero, and once again attained at a point (x_0, t_0) of C . Since (x_0, t_0) cannot lie on the level-crossing surface (assuming v is nonconstant), it must either be in the interior of C , or lie on the top surface $\{t = t_2\}$ within C . In either case, we can place a bounded cylinder within C with (x_0, t_0) in the interior or top, in violation of the maximum principle.

To show that the evolution property for level crossings yields the maximum principle, an even simpler argument suffices. For, suppose that $D \times [t_1, t_2]$ is a bounded cylinder, with a maximum (or minimum) at a point (x_0, t_0) in the interior or top of the cylinder, in violation of the maximum principle. Either way, there is a value l less than the maximum (or greater than the minimum) at (x_0, t_0) , but greater (or less than) the values of v on the sides and bottom of the cylinder. Thus, a component of the level- l crossing within the wedge $t_1 \leq t \leq t_2$ lies entirely within the interior of the cylinder, and thus does not meet the plane $\{t = t_1\}$. Thus, the evolution property fails if the maximum principle is violated. ■

As a result of the theorem, we have an evolution property for scale-space constructions more general than Laplacian-of-Gaussian filtering. For example, let D be a bounded domain, and consider scale space $T = D \times [0, \infty)$ with functions constructed as follows.

For $f \in L^\infty(D)$, solve for $u \in C(T)$,

$$\Delta u = \frac{\partial u}{\partial t} \text{ in } D \times (0, \infty)$$

$$u(x, 0) = f(x)$$

$$u(x, t) = f(x), \quad x \in \partial D, \quad t \geq 0.$$

Then let

$$v(x, t) = \frac{\partial u}{\partial t}(x, t).$$

We observe that $v = 0$ on the sides of the scale-space cylinder, and that the maximum principle holds since v satisfies the Heat Equation. Thus, level crossings of v evolve as t increases, and always meet the base $\{t = 0\}$.

Under more restrictive assumptions, one can show that the evolution property requires Laplacian-of-Gaussian scale-space construction [9]. The assumptions, however,

require that the domain be all of Euclidean space, and that the underlying equation have constant coefficients. Further, proofs of the evolution property have frequently assumed that a level-crossing component not meeting $\{t = 0\}$ will have an extremal lower point, which is an invalid restriction. Given the equivalence of the evolution property to the maximum principle, any proof that does not cite the maximum principle or essentially redo the proof is suspect. Since the proof of the maximum principle is slightly delicate, especially in the absence of strong regularity assumptions, the former course seems more appropriate.

A version of the evolution property is easy to establish when the scale space is a discrete domain. For simplicity, let us treat the case of one space dimension, and assume that we are given data f_i , $i = \dots, -1, 0, 1, \dots$. We obtain scale-space data $v_{i,k}$, where k is the scale, or level, $k = 0, 1, \dots$. The method by which v is constructed is unimportant at this point; a maximum principle states that in any rectangular grid of lattice points $\{(i, k) \mid i_1 \leq i \leq i_2, k_1 \leq k \leq k_2\}$, the maximum of $v_{i,k}$ must occur on either the bottom $k = k_1$ or the sides, $i = i_1$ or $i = i_2$. The maximum principle will surely hold true if v satisfies a construction

$$v_{i,k+1} = a_{i-1,k}v_{i-1,k} + a_{i,k}v_{i,k} + a_{i+1,k}v_{i+1,k},$$

where

$$a_{i-1,k} + a_{i,k} + a_{i+1,k} = 1 \quad \text{and all } a_{i,k} > 0.$$

Given the maximum principle for discrete data, we can easily obtain an evolution property. The evolution property will state that for any 4-connected component of grid points of $\{(i, k) \mid v_{i,k} > 0, k_1 \leq k \leq k_2\}$, this component will include points on the bottom level $k = k_1$.

B. Completeness when Gradients are Included

We now show that the zero crossings, when supplemented with gradient data along the zero crossing boundaries, are sufficient (in theory) to reconstruct the original function $f(x)$. Actually, the method is used to reconstruct $\Delta f(x)$. However, if we assume that $f(x) \rightarrow 0$ as $|x| \rightarrow \infty$, then $f(x)$ can be reconstructed by solving Poisson's equation. Details of these ideas were reported earlier in a technical report [24]. In the same report, it is observed that completeness of the representation involving gradients along the zero crossings is easily established, by a nonconstructive proof, using the Hopf maximum principle for the Heat Equation. We instead provide a constructive proof below, although it should be noted that the constructive method is unstable.

The use of gradient data for the representation also appears in [7], but the gradient data there are not limited to the zero crossings. The use of gradient data along zero crossings is discussed in [25]. Many researchers have noted from a casual observation of zero crossings of image data that zero crossings with large gradient magnitudes are of greater significance than those with low gradient magnitudes.

1) *Continuous Case:* Specifically, let Ω be a bounded connected component of $\{(x, t) \mid t \geq 0, v(x, t) \neq 0\}$, and denote by D the set $\{x \in \mathbb{R}^n \mid (x, 0) \in \Omega\}$, and by Γ the zero crossings $\partial\Omega \cap \{t > 0\}$. Let τ be a value such that $\tau > \sup\{t \mid (x, t) \in \Omega\}$. Next, denoting $\Delta f(x)$ by $g(x)$, we set $\bar{g}(x) = g(x)$ for $x \in D$, and $\bar{g}(x) = 0$ elsewhere. Let $b(x)$ be the $\bar{g}(x)$ data blurred to the level τ :

$$b(x) = \int_{\mathbb{R}^n} K(x-y, \tau)\bar{g}(y) dy$$

[see Fig. 1(c)]. Finally, for fixed y and τ define

$$w(x, t) = K(x-y, \tau-t).$$

Theorem 2: Suppose that the data $g(x)$, $x \in D$, are contained under a bounded zero crossing surface Γ of the scale-space function v . We require that $g(x)$ belongs to the class S' of tempered distributions (and thus has a Fourier transform from which g can be recovered) [26]. Then $b(x)$ can be reconstructed from the data Γ and ∇v along Γ . Thus, $g(x)$, for $x \in D$, is theoretically reconstructible.

Proof: Observe that

$$\frac{\partial w}{\partial t} + \Delta w = 0.$$

Using this fact and Green's theorem, we have

$$\begin{aligned} 0 &= \int_{\Omega} \left(\frac{\partial v}{\partial t} - \Delta v \right) \cdot w dx dt + \int_{\Omega} \left(\frac{\partial w}{\partial t} + \Delta w \right) \\ &\quad \cdot v dx dt \\ &= \int_{\partial\Omega} [(vw)\zeta + (v\nabla w - w\nabla v) \cdot n] d\sigma \end{aligned}$$

where $(n, \zeta)(x, t)$ is the outward pointing surface normal to Γ at (x, t) and $d\sigma$ is surface area measure. The gradient ∇ is with respect to spatial coordinates only. Using the facts that $(n, \zeta) = (0, -1)$ when $t = 0$ and $v = 0$ on Γ , the equation reduces to

$$\begin{aligned} &\int_D v(x, 0)w(x, 0) dx \\ &= - \int_{\Gamma} w(x, t)(\nabla v \cdot n)(x, t) d\sigma. \end{aligned}$$

Using the definitions of \bar{g} , b , and w above, the equation becomes

$$b(y) = \int_{\Gamma} K(x-y, \tau-t)(\nabla v \cdot n)(x, t) d\sigma.$$

Thus, given the zero crossing Γ and $\nabla v(x, t)$ for $(x, t) \in \Gamma$, the blurred data $b(x)$ can be constructed by a simple linear process. Since the blurring operation taking $\bar{g}(x)$ to $b(x)$ is one-to-one, $g(x)$ for $x \in D$ can theoretically be reconstructed. ■

Deblurring is, of course, a classic unstable process. The situation is not hopeless, however, since $\bar{g}(x)$ has known compact support, which might be used to advantage. and

also since errors that occur are predominantly in high-frequency components, which might not be as essential to visual interpretability. Further, one might hope that approximation methods would suffice [27]. Nonetheless, the method is provably unstable.

However, the foregoing analysis applies only to the case where Ω is bounded. Next consider an unbounded component Ω of (x, t) space with the zero crossings removed. Let τ be greater than the range of t in all of the bounded components of that space. Finally, let $\Omega_\tau = \Omega \cap \{(x, t) | 0 \leq t \leq \tau\}$. We repeat the previous argument, replacing integration over Ω by integration over Ω_τ . There is now in addition to $D = \{x \in \mathbb{R}^n | (x, 0) \in \Omega_\tau\}$, another boundary portion $D_\tau = \{x \in \mathbb{R}^n | (x, \tau) \in \Omega_\tau\}$ to consider in the surface integral. Thus, we obtain

$$\int_D v(x, 0)w(x, 0) dx + \int_{\Gamma \cap \{t < \tau\}} w(x, t)(\nabla v \cdot n)(x, t) d\sigma = \begin{cases} v(y, \tau) & (y, \tau) \in D_\tau \\ 0 & \text{otherwise.} \end{cases}$$

Here we have used the fact that $w(y, \tau) = K(x - y, 0)$ is a delta function centered at $x = y$. With $\bar{g}(x)$ and $b(x)$ defined as before, we have

$$b(y) = \int_{\Gamma \cap \{t < \tau\}} K(x - y, \tau - t)(\nabla v \cdot n)(x, t) d\sigma, \quad (x, \tau) \notin D_\tau.$$

Since $v(y, \tau)$ is unknown, the equation gives no useful information when $(y, \tau) \in D_\tau$. The missing data could be supplied by analytic continuation of $b(y)$ from $(y, \tau) \notin D_\tau$, assisted by the fact that the above equation holds with the unknown but (hopefully) small error $v(y, \tau)$ when $(y, \tau) \in D_\tau$. Once $b(y)$ has been extended in this way, $\bar{g}(x)$ is once again found by deblurring.

The lesson of this section, ultimately, is that even for bounded zero crossings supplemented with gradient data along zero crossing, reconstruction by the indicated method is still unstable.

2) *Discrete Data:* It is interesting that the analytical result above can be converted to a discrete form, without involving any discrete approximations. This is because Green's theorem can be converted to a discrete sum. For a signal or image defined on a discrete lattice, we first define a notion of a Gaussian pyramid (actually, a "monolith"), and a Laplacian pyramid, and then the concept of zero crossings in the discrete scale space. Knowledge of the zero crossings together with gradient data, in the formulation below, will translate to knowledge of the values of the scale-space function on a discrete set of points, with the points located on either side of a zero crossing. Since values are given on both sides of a zero crossing, the gradient values are implicitly encoded in the representation.

For simplicity, we treat the case of one unbounded space dimension, although the results extend easily. We are given data $f_i, i = \dots, -1, 0, 1, \dots$, and define

$$g_i = \frac{1}{4}f_{i-1} - \frac{1}{2}f_i + \frac{1}{4}f_{i+1}.$$

We define the filtered data $v_{i,k}$ recursively

$$v_{i,0} = g_i,$$

$$v_{i,k+1} = \frac{1}{4}v_{i-1,k} + \frac{1}{2}v_{i,k} + \frac{1}{4}v_{i+1,k}.$$

We also define the blurring kernel, using binomial coefficients

$$K_{i,k} = \frac{1}{4^k} \binom{2k}{i+k}, \quad -k \leq i \leq k, \quad k \geq 0,$$

and $K_{i,k} = 0$ elsewhere, for $k \geq 0$. Both v and K satisfy a discrete version of the Heat Equation, namely,

$$v_{i,k+1} - v_{i,k} = \frac{1}{4}v_{i-1,k} - \frac{1}{2}v_{i,k} + \frac{1}{4}v_{i+1,k}.$$

It is not hard to prove a discrete analog of the evolution property for zero crossings. The key, as indicated in Section II-A, is a discrete version of the maximum principle, which is easy to establish.

To formulate the reconstruction method, let Ω be a bounded 4-connected collection of pixels (i, k) with a non-empty set $D = \{i | (i, 0) \in \Omega\}$. Let T be an upper bound $T > \max \{k | (i, k) \in \Omega\}$, and define b_i to be the data $g_i, i \in D$, blurred to level T :

$$b_i = \sum_{j \in D} K_{i-j,T} \cdot g_j.$$

Finally, denote the set of pixels on the boundary of Ω by

$$\partial_{(\pm 1,0)}\Omega = \{(i, k) \in \Omega | (i \pm 1, k) \notin \Omega\},$$

$$\partial_{(0,1)}\Omega = \{(i, k) \in \Omega | (i, k + 1) \notin \Omega\},$$

$$\partial_{(0,-1)}\Omega = \{(i, k) \in \Omega | k > 0, (i, k - 1) \notin \Omega\}.$$

Then simple but messy algebra allows us to show Theorem 3.

Theorem 3:

$$4b_j = \sum_{\epsilon = -1,1} \sum_{(i,k) \in \partial_{(\epsilon,0)}\Omega} \left[\frac{v_{i,k} + v_{i+\epsilon,k}}{2} (K_{i+\epsilon-j,T-k-1} - K_{i-j,T-k-1}) - \frac{K_{i+\epsilon-j,T-k-1} + K_{i-j,T-k-1}}{2} \cdot (v_{i+\epsilon,k} - v_{i,k}) \right] + \sum_{(i,k) \in \partial_{(0,0)}\Omega} 4v_{i,k+1} \cdot K_{i-j,T-k-1} - \sum_{(i,k) \in \partial_{(0,0)}\Omega} 4v_{i,k} \cdot K_{i-j,T-k}.$$

To reconstruct data by the above equation, choose a connected component of $\{(i, k) | v(i, k) > 0\}$ (or, respectively, < 0). If the component extends to infinity in

either coordinate, truncate the domain to become a convenient bounded collection of pixels, and denote the result by Ω . We store the sets $\partial_{(\pm 1,0)}\Omega$, $\partial_{(0,\pm 1)}\Omega$, and D as defined above. For pixels (i, k) in $\partial_{(1,0)}\Omega$ (respectively, $\partial_{(-1,0)}\Omega$), we store the information $v_{i,k}$ and $v_{i+1,k}$ (respectively, $v_{i,k}$ and $v_{i-1,k}$). For pixels (i, k) in $\partial_{(0,1)}\Omega$, we store the data $v_{i,k+1}$ and for $\partial_{(0,-1)}\Omega$ pixels we store $v_{i,k}$. Using the above equation, we choose a T and reconstruct the blurred data b_j . To reconstruct the data g_i for $i \in D$, it suffices to deblur the b_i data by solving for g_i in the linear equations defining b_i . In fact, the system is overdetermined, although still poorly conditioned, especially if $|D|$ or T is large.

In order to make the computations feasible, it is necessary to modify the formulas for a bounded spatial domain. For example, a common approach to constructing a discrete scale-space function involves solving a bounded domain problem, with $-N \leq i \leq N$, by setting $v_{i,k} = 0$ for $i = \pm N$. The blurring kernel K is changed by this modification, but the proposition carries over with little change.

There is another version of the formula for b_j above, and thus for the reconstruction of the image data, where the stored information consists of the precise locations of the zero crossings, assuming a linear interpolation of values between grid points. The gradient data, in the form of the difference of the neighboring grid values, still need to be included. With this information, the formulas simplify some, but the information in the representation is clearly equivalent, so we omit the details.

III. COMPLETENESS AND A BRIEF SURVEY OF RECONSTRUCTIONS

Under certain restrictions on the class of signals, it can be shown that the zero crossings in scale space form a complete representation of the signal data. There have been many different studies and theorems along these lines for applications to computer vision. An oft-cited study, for the case of signals in one dimension, is provided by Logan [28]. Logan's theorem requires that the function be bandpass, and satisfy certain technical conditions concerning the signal's Hilbert transform. Further, the result is intended for the reconstruction of a function represented by the zero crossings at exactly one level of resolution. While feasible as a complex analytic result, our interest here is in the more practical possibility of reconstruction given zero crossings at multiple scales of resolution in scale space.

More complex analytic results can be used when the image data are polynomial. In fact, when $f(x)$ and thus $\Delta f(x)$ is a polynomial in $x \in \mathbb{R}^n$, then $v(x, t)$ is a polynomial in $(x, t) \in \mathbb{R}^{n+1}$. Accordingly, the zero crossings are part of the *analytic varieties* of the polynomial v as studied in algebraic geometry. It is well known that the varieties in C^n determine the complex polynomial defined on n complex variables. It is not as commonly used, but nonetheless true, that an n -dimensional subportion of the intersection of the analytic variety with \mathbb{R}^{n+1} determines

an irreducible polynomial. A proof of this result, supplied by Mumford, is given in the Appendix. Related proofs are given by Huang and Sanz [29] and Curtis, Oppenheim, and Lim [30]. Thus, the case of polynomial data can be settled with algebraic geometry.

However, these results depend heavily on the assumption that the data are polynomial. It is obvious that there are many different functions, even smooth functions, that have identical zero crossings. Interestingly, the completeness in the case of polynomial data applies equally well to the scale-space function $v(x, t)$ as to the original data $f(x)$ or, for that matter, any single level of the scale-space function $v(x, t_0)$. That is, the polynomial $f(x)$ is determined by its zero crossings (providing it is irreducible as a polynomial), a level of the scale-space function $v(x, t_0)$ is determined by its zero crossings, providing it is irreducible, or the scale-space function $v(x, t)$ is determined by its zero crossings, providing it is irreducible. In any of the above cases, $f(x)$ can theoretically be reconstructed. However, it is in some sense more likely that $v(x, t)$ will be irreducible, and we expect the reconstruction of f from the function v to be more stable than reconstruction from a single level of v , or from the zero-crossing curves of f alone.

Yuille and Poggio [31] prove a similar result: showing that when $\Delta f(x)$ and hence $f(x)$ is polynomial in x , and if $n = 1$, then reconstruction from zero crossings is theoretically possible. They refer to the validity of the observation for larger n . Their method relies on an expansion of the analytic structure of a zero-crossing contour, using arbitrarily high derivatives of the contour, at two points of the zero crossing in scale space. While the method is constructive, it is not likely to lead to stable reconstructions.

Curtis, Shitz, and Oppenheim [32] and Sanz and Huang [29] extend the above results to the case when $f(x)$ is a band-limited function which is irreducible as an entire function. These results are likewise of an algebraic geometry nature, and rely on the fact that there is a finite Fourier series of the band-limited function that provides a polynomial of several complex variables for representing f .

Since the determination of a polynomial by its varieties is essentially an analytic continuation result, stability of the reconstruction is unlikely. That is, small errors in measurement of the zero crossings could lead to arbitrarily large errors in the determination of $\Delta f(x)$. Put differently, there can be widely different initial data leading to nearly identical zero-crossing data.

Worse, settling the case for polynomial data, or even irreducible entire band-limited data, says little about the general case of continuous initial data. Although the Stone-Weierstrass theorem says that a continuous function can be uniformly approximated by a polynomial on a compact set, the zero crossings depend on the initial data globally, and the dependence cannot be localized. Further, the lack of stability means that the approximation is irrelevant. The situation is similar to the fact that a poly-

nomial of a single variable with all real roots is determined by its zeros, but that given all the zeros of a continuous function, one knows nothing more than the zeros.

Based on these results of completeness under certain restrictions, there have been a number of attempts at reconstructions from zero crossings. The hope is that natural image data will sufficiently restrict the class of admissible functions that some stable reconstruction method can be found. We briefly survey below three such attempts. In Section IV, we provide a formulation for a new attempt at reconstruction.

A. Curtis–Oppenheim

Curtis and Oppenheim, using results that show that band-limited images are completely represented by their zero crossings under certain conditions, present some experiments with reconstruction [33]. In their formulation, they begin with a band-limited image $f(x, y)$, and record a large number of locations in the (x, y) domain where f crosses some threshold (such as zero, assuming that f has both positive and negative regions). These locations are recorded with great precision, to ultra-subpixel accuracy. The image is reconstructed from this information. Thus, in their work, the image is reconstructed from a number of level-crossing contours. Clearly, the assumption that the image is band-limited is crucial, for otherwise there are an infinity of images having precisely the same level crossings. In essence, the work constitutes reconstruction from thresholded imagery, but with the understanding that the image is band-limited, and that the level-crossing contours are recorded with great accuracy.

The method begins by writing the discrete image f as the inverse Fourier transform of its discrete Fourier transform F :

$$f(k, l) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i u k / N} e^{2\pi i v l / N}.$$

This function is interpolated to an entire band-limited periodic function, $f(x, y)$, given by

$$f(x, y) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i u x / N} e^{2\pi i v y / N}.$$

It is the locations of zeros of $f(x, y)$ that are recorded, say (x_n, y_n) , $n = 1, \dots, K$. We can then write the K equations

$$\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i u x_n / N} e^{2\pi i v y_n / N} = 0,$$

$$n = 1, 2, \dots, K.$$

If the mean value μ of $f(x, y)$ is not zero, then the additional equation $F(0, 0) = \mu N^2$ is included. Otherwise, one more equation is needed, giving the location and value of $f(x, y)$ at a single location where f is nonzero. Assuming the former case, we have K linear equations in the unknowns $F(u, v)$, $(u, v) \neq (0, 0)$, where (u, v) ranges over the known (band-limited) support of F . By using many more equations than unknowns, and computing the

least squares solution, reasonable reconstructions are obtained. However, in the examples shown, the number of independent nonzero spectral components is roughly 200. Further, the locations of these spectral components were chosen carefully, in advance, and the image to be reconstructed was obtained from a full-resolution image by setting all except the 200 Fourier coefficients to zero. By tiling Fourier space (as Rotem and Zeevi do; see below), a general image could be reconstructed by this method using the zero crossings of multiple images, each obtained by filtering the initial image. However, since a 256-by-256 image contains 64K independent spectral components, it is not clear how many levels will be needed. Nonetheless, the method provides a theoretically possible means for representation and reconstruction from multiple images of zero crossings.

B. Rotem–Zeevi

Rotem and Zeevi, in a paper in which they extend the Logan theorem for application to two-dimensional images, experiment with reconstructions from zero-crossing data [34]. Their method is based on first decomposing the given image $f(x, y)$ into a sum of images $f_k(x, y)$, with each f_k obtained from f by a filtering operation in Fourier space. Each f_k has the property that its discrete Fourier transform has nonzero coefficients in a region that is bandpass of one octave or less in at least one of the two spectral dimensions. Further, the low-frequency components are omitted from the reconstruction process, and simply stored as part of the representation, and an attempt is made to reconstruct only a low-pass version of f , where all frequencies above half of the maximum are omitted. Thus, for a 256-by-256 image, where spectral frequencies range from -128 to 127 , only the frequencies in the range -64 to 64 are included in the reconstruction, and the spectral frequencies where both components u and v satisfy $|u| \leq 8$, $|v| \leq 8$, are stored separately, and excluded from the reconstruction. The remaining spectral domain is decomposed into 19 regions, so that there result 19 images, f_k , $k = 1, \dots, 19$. If f_0 denotes the low-pass image composed of frequencies below 8, and if $g(x, y)$ is the high-pass image composed of frequencies above 64 (in either dimension), then

$$f = f_0 + \sum_{k=0}^{19} f_k + g.$$

By keeping the filters yielding the decomposition symmetric (i.e., including the component for $(-u, -v)$ if the coefficient for (u, v) is included), we can be sure that if f is a real-valued image, then each of the images in the decomposition will be real valued.

Rotem and Zeevi reconstruct each of the f_k from the information $\text{sgn}(f_k)$, $k = 1, \dots, 19$. Since each f_k is a bandpass of f , they correspond, crudely, to the levels of the scale-space function v that we use to reconstruct f . However, the analogy is poor, since the f_k are true bandpass images, with one octave or less in frequencies in at least one of the two spectral dimensions, whereas a level

of v is only an approximate isotropic bandpass filter of f . Further, the number of free variables in any given f_k is small—the greatest number of spectral components used in any of the f_k is 1024, whereas the $\text{sgn}(f_k)$ have 64K bits of information, for each k . In any case, the reconstruction of an f_k works as follows.

Suppose that f_k is bandpass in the horizontal spectral dimension. It is then easily shown that each row of f_k is a one-dimensional bandpass signal. (Alternatively, if f_k is bandpass in the vertical spectral dimension, then columns of f_k are bandpass.) The spectral support of the rows or columns is the corresponding band in the horizontal or vertical dimension of the spectral domain of f_k . The reconstruction begins by finding individual rows or columns that can be reconstructed from the sign data along the row or column. For a one-dimensional row S which is bandpass, denote by B_p the bandpass filter; thus, $S = B_p(S)$. Given the data $\text{sgn}(S)$, let $S_0 = B_p(\text{sgn}(S))$. The algorithm to recover S proceeds iteratively, beginning with S_0 , and iteratively setting

$$S_{n+1} = S_n - c[B_p(\text{sgn}(S_n)) - S_0].$$

The iteration is not guaranteed to converge, but if it does, one generally has that $\text{sgn}(S_\infty) = \text{sgn}(S)$. Providing all the conditions for uniqueness of the representation of the one-dimensional signal by its zero crossing are met (i.e., Logan's conditions), then this will imply that S_∞ and S differ only by a multiplicative constant. Having reconstructed a number of rows (or columns), modulo multiplicative constants, the one-dimensional signals must be individually scaled. By reconstructing a one-dimensional signal in a transverse direction, ratios can be obtained to scale the individual rows (or columns) modulo a single multiplicative constant, relative to the whole image. If the image is bandpass in the vertical dimension, then columns can be reconstructed. If the image is only low-pass in this dimension, Zeevi and Rotem show that it will be bandpass in a diagonal direction, and so some diagonal can be reconstructed. In practice, they reconstruct several transverse one-dimensional signals, and use averages to best scale the individual rows. Reconstruction along columns that do not initially converge can be assisted by initialization with data obtained from diagonals or rows, and from horizontally blurring the partially reconstructed image.

Once each f_k is reconstructed to within a multiplicative scale factor, a scale factor is applied, based on the assumption that the variance of each f_k is known in advance. The final reconstruction of f is performed using the reconstructed f_k , the stored f_0 , and by omitting g . The result shown in their paper is visually good, and they show that errors are mostly concentrated on the edges of the original image.

C. Sanz-Huang

More recently, Sanz and Huang have shown examples of reconstructions of images using zero crossings of the scale-space function recorded at four levels [29].

First, four levels of the scale-space function $v(x, y, t_i)$ are chosen, $i = 1, 2, 3, 4$, where v is the scale-space function of the initial image f . The information that is stored in the representation is $\text{sgn}(v(x, y, t_i))$ for $i = 1, \dots, 4$. The operator that takes $f(x, y)$ to $\Sigma v(x, y, t_i)$ is linear, and a pseudoinverse is computed. This operator, which cannot restore the very low spectral components or the very high spectral components, is nonetheless a convolution operator (assuming an unbounded domain), given by convolution against a function ϕ . That is,

$$f \approx \phi * \sum_i v(\cdot, \cdot, t_i).$$

The reconstruction algorithm proceeds iteratively, beginning with

$$S_0 = \phi * \sum_i \text{sgn}(v(\cdot, \cdot, t_i)).$$

Given S_n , the next iteration is obtained by first computing the scale-space levels of the estimate S_n

$$v_n(\cdot, \cdot, t_i) = \Delta K(\cdot, \cdot, t_i) * S_n,$$

and then modifying v_n so that the data have the right sign

$$\begin{aligned} \bar{v}_n(x, y, t_i) &= \begin{cases} |v_n(x, y, t_i)| & \text{if } \text{sgn}(v(x, y, t_i)) > 0 \\ -|v_n(x, y, t_i)| & \text{if } \text{sgn}(v(x, y, t_i)) < 0. \end{cases} \end{aligned}$$

Finally, the \bar{v} 's are summed and filtered to obtain the next estimate

$$S_{n+1} = \phi * \sum_i \bar{v}_n(\cdot, \cdot, t_i).$$

This procedure is not guaranteed to converge, but was used for experimental purposes anyway. When applied to large, natural grayscale images, they resulted in reconstructions that had scale-space functions whose signs agreed, on the four levels, with the original sign data of the representation, at roughly 80 percent of the pixels. However, the reconstructions, while clearly related to the originals, had a rather stylized or "impressionistic" appearance, and many perceptual differences from the originals.

D. Comments

Each of the above reconstruction techniques attempts to find a function that satisfies the given zero-crossing information by taking some measurement of the error of the zero crossings of the reconstruction with the desired zero crossings. We might call this generic approach one of minimizing data error: the error in realizing the zero-crossing data. As yet another version of minimizing data error, we could take as a measure of the error the square integral of the L^2 difference between the sign of the scale-space function and the desired sign of the scale-space function:

$$\int_0^\infty \|\text{sgn}(v(\cdot, \cdot, t)) - \text{sgn}(v_0(\cdot, \cdot, t))\|^2 dt,$$

where $\text{sgn}(v_0)$ is the given data in scale space, and $v = \Delta K * f$. This functional is not differentiable with respect to the reconstructed image f due to the discontinuous signum function. However, if we replace the signum function sgn with a C^∞ approximate, say ψ , then the error measure becomes a smooth function of the f data. Suppose that instead of being given the sign of the scale-space data, we are instead given $S_0(x, y, t) = \psi(v_0(x, y, t))$. Suppose that f is the current estimate of the original image. Then the gradient of the error measure with respect to the estimate f is proportional to

$$\int_0^\infty [(\psi(v(\cdot, \cdot, t)) - S_0(\cdot, \cdot, t)) \cdot \psi'(v(\cdot, \cdot, t)) * \Delta K](x, y, t) dt$$

where $v = \Delta K * f$. Thus, a simple gradient descent procedure can be formulated to minimize the error. The result is a particularly simple linear iterative process. Of course, the representation is now the approximate signum of the v_0 data, $\psi(v_0)$, rather than the simple boolean information about the sign of the filtered data or the location of the zero crossings. Further, convergence is not guaranteed and can be very slow. If the true signum data are used as the target data in place of S_0 , then convergence of the iterative process is much worse.

Instead of pursuing this approach, we instead develop the idea of minimizing equation error, discussed in the next section.

IV. FORMULATION AND METHODS

A. Equation Error Minimization

In this section, we describe a method for reconstructing an image from its zero crossings based on minimizing equation error. The method was reported for a different application (deblurring) in conference proceedings [35], and preliminary results on reconstructions from zero crossings were shown at a different conference [36]. The use of minimization of equation error to solve ill-conditioned inverse problems is motivated by a related technique that has been applied to identification problems for elliptic equations [37], such as in porous-media flow problems (i.e., oil recovery). In these applications, the method proved more favorable than more typical identification problem methods, related to minimization of data error. Although one can provide heuristic arguments motivating the selection of equation error minimization for inverse problems, a thorough mathematical justification of the stability and advantages of the method is still lacking.

Our approach to reconstruction from zero crossings is to reconstruct the function v in scale space, and make use of the fact that v satisfies the Heat Equation. The discretized data can be viewed as a multilevel grid of units communicating locally. The essence of the idea is that the units should achieve values satisfying the given zero-crossing constraints, and also satisfying, to the extent possible, a discrete version of the Heat Equation. As in a

network, the values are updated iteratively using information from local values to minimize a measure of error. As such, the method is amenable to implementation using networks of analog processors connected in a mesh by tunable resistive links.

We first sketch the formulation of the approach in a continuous domain. Consider the initial image $f(x)$ and its scale-space Laplacian-of-Gaussian filtered data $v_0(x, t)$. We assume that the zero crossings in scale space of v_0 are given, in the form of knowledge of $\text{sgn}[v_0(x, t)]$ at all points. We then pose the problem:

$$\text{Find } v \text{ minimizing } \left\| \Delta v - \frac{\partial v}{\partial t} \right\|^2,$$

$$\text{subject to } \text{sgn}[v(x, t)] = \text{sgn}[v_0(x, t)].$$

Here, the norm $\| \cdot \|$ is the L^2 norm over the scale space $\mathbb{R}^n \times \mathbb{R}^+$.

We in fact advocate a slightly different measure of the equation error. We transform the Heat Equation, a second-order partial differential equation, into a first-order system

$$\begin{aligned} \nabla v &= \sigma, \\ \nabla \cdot \sigma &= \frac{\partial v}{\partial t}. \end{aligned}$$

The problem is now reformulated as the search for a pair of functions (v, σ) minimizing

$$\left\| \nabla \cdot \sigma - \frac{\partial v}{\partial t} \right\|^2 + \|\nabla v - \sigma\|^2$$

subject to

$$\text{sgn}[v(x, t)] = \text{sgn}[v_0(x, t)]$$

These problems can be given a more precise formulation using the theory of variational inequalities [38]. The spaces of admissible functions (functions satisfying the inequality constraints) for the unknowns v and σ should be defined as convex subsets of appropriate Sobolev spaces. When this is done, the problem becomes a standard "obstacle problem," for which existence and uniqueness of a solution can easily be shown. Further, the only local minimum of the functional is also the global minimum. However, we can only be sure that the Heat Equation is satisfied if the error is actually zero. Accordingly, we know that if we minimize the equation error, we (at least in theory) will have no problems with local minima, and will find a single global minimum; further, we know that we can always find a global minimum, but that we have a solution to our reconstruction problem (that is, a solution to the Heat Equation with the given constraints) only if the error at that minimum is zero. In general, existence of a solution to the Heat Equation with given sign constraints is not guaranteed.

Numerical methods for the solution of constrained quadratic optimization problems, such as this one, are

much studied in numerical analysis. Related optimization problems frequently occur in computer vision [39]. In the next two sections, we discuss some of the implementation details for the experiments described in Section V, concentrating on the discrete formulation peculiar to the equation error minimization problem for reconstructions from zero crossings.

B. Discrete Formulation

We now describe the discrete formulation of these ideas. The case of a finite-difference discretization for $n = 1$ will be described in some detail. Extension to two spatial dimensions involves some complications which will be summarized more briefly. A finite-element formulation is also given for the case $n = 1$.

1) *One Spatial Dimension*: From the given initial data f_i , we recursively construct the scale-space function $v_{i,k}$ from the formulas:

$$v_{i,0} = \frac{1}{4}[f_{i-1} - 2f_i + f_{i+1}]$$

$$v_{i,k+1} = \frac{1}{4}[v_{i-1,k} + 2v_{i,k} + v_{i+1,k}].$$

To make the computation feasible, the space and scale domains must be bounded. Various approaches to handling the spatial border points are possible. A particularly convenient one, analogous to Neumann boundary conditions, is defined as follows. Let us assume that f_i is defined for $0 \leq i \leq N-1$. We then extend f_i to be periodic with period N , according to the formulas

$$f_{N+i} = f_{N-1-i}, \quad i = 0, 1, \dots, N-1,$$

$$f_{2N+i} = f_i, \quad \text{all } i.$$

Then the values $v_{i,k}$ are defined for all i and all $k \geq 0$. The scale-space function $v_{i,k}$ is periodic with period $2N$ in the variable i , and has the same structure as the extended f_i data. Specifically, the data from $i = 0, \dots, N-1$ are repeated in reverse order from $i = N, \dots, 2N-1$, and then extended as a $2N$ block periodically (see Fig. 2).

The v data approximate the Laplacian applied to the Gaussian-blurred image data. Associated with the computed scale-space data $v_{i,k}$ are the sign data $S_{i,k}$, which are defined to be 1, 0, or -1 according to whether $v_{i,k}$ is positive, zero, or negative, respectively. In practice, the v data and resulting S data are calculated only for $0 \leq i \leq N-1$, and for $0 \leq k \leq T$, for some arbitrary cutoff T .

Having computed the representation in the form of the $S_{i,k}$ data, we proceed to the reconstruction. The discrete reformulation of the Heat Equation as a first-order system involves the introduction of a second grid of data, $\sigma_{i,k}$ which ideally satisfies

$$\sigma_{i,k} = \frac{1}{2}[v_{i,k} - v_{i-1,k}].$$

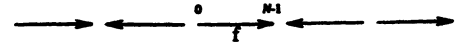


Fig. 2. Borders are handled by extending the original data f . The data f_0, \dots, f_{N-1} are repeated alternately from left to right, and then from right to left, to form a doubly infinite signal that is periodic with period $2N$. Note that the values at the endpoints, e.g., the datum f_{N-1} , appears twice in succession.

Then the equation error is defined by the quadratic functional

$$E = \sum_{k \geq 0} \sum_{i=0}^{N-1} \left(\frac{v_{i,k} - v_{i-1,k}}{2} - \sigma_{i,k} \right)^2 + \left(v_{i,k+1} - v_{i,k} - \frac{\sigma_{i+1,k} - \sigma_{i,k}}{2} \right)^2.$$

We minimize E subject to the constraints

$$v_{i,k} \geq 0 \quad \text{when } S_{i,k} = 1$$

$$v_{i,k} \leq 0 \quad \text{when } S_{i,k} = -1.$$

A further constraint is needed to ensure that the solution $v \equiv 0$ is not obtained. In principle, this could be done by specifying a single value for v at some node. In practice, such a constraint would propagate only very slowly through the grid, making convergence difficult. We chose instead to specify v at the maximum level $k = T$, where it is ordinarily quite smooth. Little is added to the representation by inclusion of the $v_{i,T}$ data, since these data could be quite accurately represented by a very few floating-point values specifying the low-order discrete Fourier coefficients. Recall that a similar approach is utilized by Zeevi and Rotem (Section III-B). (In experiments involving fixed gradient data along the zero crossings, the zero solution is no longer admissible, and the $v_{i,T}$ data are not included in the representation.)

As in the continuous case, the minimization of the (coercive), quadratic functional on a convex set results in a unique local minimum which is also the global minimum. The difficulty comes about because of the large number of variables. Without care, convergence can be extremely slow. Further, we do not know in advance how "shallow" the minimum will be on the convex set. When the minimum is shallow, then the problem is unstable, and we expect to see dependence on initial conditions in the iterative minimization algorithms. Such is the case for the experiments with reconstructions from zero crossings alone. When the minimum is sharper, convergence should be faster and independent of the initial conditions, as is the case in the experiments with reconstructions from zero crossings with gradient information.

In the experiments described in Section V, we used a conjugate-gradient approach to minimizing the error functional. The inequality constraints were imposed by means of a penalization method, adding to E a term that grows quadratically in the variable whenever a constraint is violated. By placing a large weight on the penalization term, a functional suitable for minimizing is obtained that will

ensure, upon convergence, that the constraints hold approximately. Specifically, we use a penalization term of the form

$$\lambda \cdot \frac{1}{2} \sum_{i,k} |\text{sgn}(v_{i,k}) - S_{i,k}| \cdot v_{i,k}^2.$$

The gradient of the equation error with respect to the unknowns can be represented as a sum of discrete convolutions applied to the data v and σ , which may be regarded as image arrays. Then for E equal to the equation error as above, the values of $\partial E / \partial v_{i,k}$ and $\partial E / \partial \sigma_{i,k}$ form image arrays also, which we will denote by $\nabla_v E$ and $\nabla_\sigma E$. The formulas (valid for $k > 0$) are

$$\begin{aligned} \nabla_v E &= \begin{bmatrix} 0 & -4 & 0 \\ -1 & 10 & -1 \\ 0 & -4 & 0 \end{bmatrix} * v + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -4 & 4 \\ 0 & 2 & -2 \end{bmatrix} * \sigma, \\ \nabla_\sigma E &= \begin{bmatrix} -2 & 2 & 0 \\ 4 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix} * v + \begin{bmatrix} 0 & 0 & 0 \\ -1 & 6 & -1 \\ 0 & 0 & 0 \end{bmatrix} * \sigma. \end{aligned}$$

Here, the three-by-three arrays are masks that are employed in three-by-three local convolutions against the array data v and σ . The boundaries can be conveniently handled by placing borders one pixel wide around the data arrays. These border pixels are set so as to satisfy the chosen boundary conditions. The sign penalty adds a term

$$\lambda |\text{sgn}(v) - S| \cdot v$$

to $\nabla_v E$.

The minimization by conjugate gradients [40] is implemented as follows. Observe that the equation error E as formulated above is a quadratic functional of the unknowns v and σ . Combining the unknowns v and σ into a single vector, say x , the equation error is thus equivalent to the quadratic functional

$$E = \frac{1}{2} x^T A x - h^T x + E_0$$

for some positive definite matrix A , vector h , and scalar constant E_0 . Adding the penalization terms to E maintains the same basic form, although the matrix A becomes a function of v . Nonzero values for the quantities h and E_0 arise from fixed grid values on the boundaries. The method of conjugate gradients is ordinarily formulated as the minimization of a functional of exactly this form. The computation involves forming the inner products of certain vectors with each other, and products of vectors with the matrix A . However, there is no need to find A , h , and E_0 explicitly. Instead, we will make use of the fact that their products with relevant vectors have known forms. Observe that

$$\nabla E = Ax - h.$$

Hence, if p is an arbitrary vector (e.g., one of the conjugate gradients), the product Ap can be computed by

evaluating the expression of ∇E , according to the convolutions given above, with p substituted for x , which is the combined v and σ data, and h set to zero. Setting h to zero is achieved simply by replacing all constants on the boundaries by zeros. This also has the effect of setting E_0 to zero. The other matrix product of interest, $p^T Ap$, is computed by evaluating directly the expression for E , again substituting p for x and setting boundary constants to zero. Note that if reflected boundary conditions are in effect, then values for p must be reflected in the same way. Thus, we see that all the matrix products necessary for a conjugate gradient calculation can be computed without knowing A explicitly.

It is worth noting that this formulation of the conjugate gradient minimization is readily adapted to heterogeneous constraints according to which v or σ are held fixed at arbitrary locations in the grid. One simply sets the values of p to 0 at the corresponding positions. In the iterative updating, those points are skipped over, although they enter into the calculation of the equation error and the matrix products.

2) *Two Spatial Dimensions:* We now turn to the finite-difference discretization for the case $n = 2$. We define the scale-space grid as $v_{i,j,k}$, $k \geq 0$, and formulate the Heat Equation as

$$\begin{aligned} v_{i,j,k+1} &= \frac{1}{16} [v_{i-1,j-1,k} + 2v_{i-1,j,k} + v_{i-1,j+1,k} \\ &\quad + 2v_{i,j-1,k} + 4v_{i,j,k} + 2v_{i,j+1,k} \\ &\quad + v_{i+1,j-1,k} + 2v_{i+1,j,k} + v_{i+1,j+1,k}] \end{aligned}$$

(2-D Discrete Heat Equation)

Reformulation of this equation as a first-order system is not so straightforward as in the one-dimensional case. A fairly natural approach proceeds with introduction of variables $\sigma_{i,j,k}^{(1)}$, $\sigma_{i,j,k}^{(2)}$, and $w_{i,j,k}$ which ideally satisfy

$$\sigma_{i,j,k}^{(1)} = \frac{v_{i,j,k} - v_{i-1,j,k}}{2},$$

$$\sigma_{i,j,k}^{(2)} = \frac{v_{i,j,k} - v_{i,j-1,k}}{2},$$

$$w_{i,j,k} - v_{i,j,k} = \frac{\sigma_{i+1,j,k}^{(1)} - \sigma_{i,j,k}^{(1)}}{2},$$

and

$$v_{i,j,k+1} - w_{i,j,k} = \frac{\sigma_{i,j+1,k}^{(2)} - \sigma_{i,j,k}^{(2)}}{2}.$$

When these equations hold exactly, the data v are blurred from level k to level $k + 1$ by convolution against the two-dimensional discrete Heat Equation. Analogously to the $n = 1$ case, the equations above can be converted to a quadratic measure of equation error in the unknown variables v , w , $\sigma^{(1)}$, and $\sigma^{(2)}$. The sign constraints on v can again be implemented using quadratic penalty terms.

The resulting error functional can then be minimized by a conjugate-gradient procedure in the same way as described for $n = 1$. It is worth noting that this approach results in equations for the gradient of the equation error which are linear in the unknowns, and that a component of the gradient with respect to an unknown at (i, j, k) depends only on values at neighboring grid points. Thus, this formulation is well suited to a hardware implementation involving only small-kernel convolutions and (for the conjugate-gradient method) accumulation of simple global sums.

3) *A Finite Element Approach:* The discretization of the Heat Equation can alternatively proceed by a finite-element approach [41]. For the case $n = 1$, we assume a regular rectangular discrete grid, $(i \cdot h, k \cdot s)$, $i = 0, \dots, N - 1$ and $k = 0, 1, \dots$. Here h and s specify the scaling of the two axes relative to each other. (In the finite-difference method, we implicitly set $h = 1/2$, $s = 1$.) The simplest elements can be formed by splitting each grid square into two triangular regions, along one of the two chosen diagonals (standard triangular elements); see Fig. 3. The unknown functions, v and σ , can be represented as continuous piecewise linear functions, linear within each triangular region, by simply specifying the value of the functions at each grid point. Accordingly, the unknowns are the values of $v(ih, ks)$ and $\sigma(ih, ks)$, for all i and k , and will be denoted by $v_{i,k}$ and $\sigma_{i,k}$. The simple piecewise linear elements suffice since the equation was converted to a first-order system. It is then a simple (albeit tedious) matter to compute the equation error for such functions in terms of the unknowns. Like the finite-difference formulation, it turns out that the equation error is a quadratic function in the unknowns, and that there is locality of dependence.

The case $h = 1/2$, $s = 1$ is very common, and so we will give the gradient equations for this case. As was done for the finite-difference case, the result can be represented as a sum of discrete convolutions, with the data v and σ considered as image arrays. Once again, except near border pixels, these gradient values $\partial E/\partial v_{i,k}$ and $\partial E/\partial \sigma_{i,k}$ can be viewed as images obtained by convolutions against the grid arrays $v_{i,k}$ and $\sigma_{i,k}$. The formulas are

$$\nabla_v E = \begin{bmatrix} 0 & -4 & 0 \\ -1 & 10 & -1 \\ 0 & -4 & 0 \end{bmatrix} * v + \frac{1}{3} \begin{bmatrix} -4 & 4 & 0 \\ 1 & -6 & 5 \\ 0 & 2 & -2 \end{bmatrix} * \sigma,$$

and

$$\nabla_\sigma E = \frac{1}{3} \begin{bmatrix} -2 & 2 & 0 \\ 5 & -6 & 1 \\ 0 & 4 & -4 \end{bmatrix} * v + \frac{1}{3} \begin{bmatrix} 1 & 1 & 0 \\ -2 & 12 & -2 \\ 0 & 1 & 1 \end{bmatrix} * \sigma.$$

Separate equations are needed for $k = 0$ ($\partial E/\partial v_{i,0}$ and $\partial E/\partial \sigma_{i,0}$) and on the borders. Note that a certain amount of asymmetry results simply from choosing triangular elements based on one of two possible diagonal cuts through each rectangular grid. The sign penalty adds the same term to $\nabla_v E$ as for the finite-difference formulation.

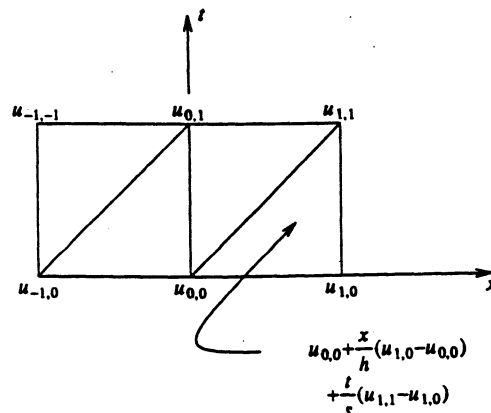


Fig. 3. Triangular, piecewise-linear, finite elements.

C. Image Reconstruction from Scale-Space Data

Once the approximate scale-space function \tilde{v} has been reconstructed by the above minimization process, there remains the problem of obtaining the image \tilde{f} to which it corresponds. First, the discrete reconstructed data are summed over the scale dimension to give

$$\tilde{r} = - \sum_{k=0}^T \tilde{v}_k.$$

The \tilde{r} data can be seen to be the discrete approximation to $f - u_{T+1}$, where the u data are defined in terms of convolution of f against a discrete Gaussian kernel, $K_{i,k}$, as defined in Section II-B-2. The problem of computing the approximation \tilde{f} to f thus reduces to inverting $I - G_{T+1}$, where G_{T+1} is the operator yielding $T + 1$ levels of Gaussian blur. That is, $(I - G_{T+1})\tilde{f} = \tilde{r}$. The simplest and cheapest method is simply to incorporate the u_{T+1} data into the representation, so that an approximation to the original f can be obtained by computing \tilde{r} and adding u_{T+1} . Since the u_{T+1} data are quite blurry, including them does not add much information to the representation.

However, since the operator $I - G_{T+1}$ is well behaved, it is possible to reconstruct \tilde{f} modulo an additive constant and a multiplicative constant, without inclusion of the u_{T+1} data in the representation, by computing the pseudoinverse of $I - G_{T+1}$ using a singular-value decomposition (SVD). The resulting \tilde{f} always has an average value of zero; therefore, the average value of the original image is added to \tilde{f} to complete the reconstruction. Unfortunately, computing the SVD can be prohibitively expensive for images larger than about 16 by 16.

Providing the operator $I - G_{T+1}$ can be viewed as a convolution, the singular-value decomposition can be computed efficiently, even for large images, using the discrete Fourier transform of $I - G_{T+1}$ or, equivalently, the inversion may be done by means of the discrete convolution theorem. Depending upon the method of handling borders, $I - G_{T+1}$ may or may not be a convolution. If a Fourier method is applied when the borders are handled in such a way that the operator is not a convolution, then even though border effects are confined to a neighborhood

of a few pixels from the border, the inversion yields poor results. However, some important border methods can be viewed as convolutions. In particular, for the Neumann boundary conditions model, as discussed in the previous section, the operator is equivalent to a convolution operation against a doubly periodic image that is twice the size of the original in each dimension, obtained by reflecting the image in each dimension. When the Fourier method is used to invert the operator $I - G_{T+1}$ computed using the doubly reflected periodic image, the results are excellent.

The inversion is complicated by the singularity that arises out of the additive-constant ambiguity: the "DC-component" of $I - G_{T+1}$ has zero amplitude. Further, the Fourier coefficients belonging to low frequencies will be small in value. A cutoff must be chosen judiciously in order to avoid amplifying low-frequency errors in the inversion process.

Whether the operator $I - G_{T+1}$ is a convolution or not, provided it is linear, and provided all eigenvalues of G are strictly less than one, the inversion may be obtained from the series expansion

$$(I - G_{T+1})^{-1} = \sum_{m=0}^{\infty} (G_{T+1})^m,$$

which can usually be computed quite easily, if somewhat intensively.

The results obtained by using the appropriate method for the given boundary conditions are, in all cases, very similar to those obtained using the simpler method by adding u_{T+1} to \bar{f} as described earlier, assuming that u_{T+1} is part of the representation. Consequently, the reconstructions shown below were performed by the latter method.

V. RESULTS

A. One-Dimensional Experiments

Our first experiments were done with one-dimensional data taken from a scanline of a digitized image. Fig. 4 shows a plot of the original data. The Laplacian monolith v_0 of the original image was computed, as described in Section IV-B-1, to a depth of 40 levels of blurring by the "1 2 1" mask. Fig. 5 shows the sign of the function v_0 , using black for negative regions, white for positive regions, and gray for pixels whose values lie within machine-precision zero. Reconstruction proceeded by minimization of equation error, using conjugate gradients with a penalty term for values of v which differ in sign from the corresponding points of v_0 , and holding the topmost (most blurred) level fixed equal to the corresponding level of v_0 . The methods for the calculation of the gradients and of the construction were described in Sections IV-B-1 and IV-C. The initial estimate for v was $\text{sgn}(v_0)$, and the initial estimate for σ was the gradient of this initial v .

The computation was run for over 40 000 iterations, achieving an equation error plus penalty of less than 10^{-10} with only 77 sign violations out of 5120 grid points. (It

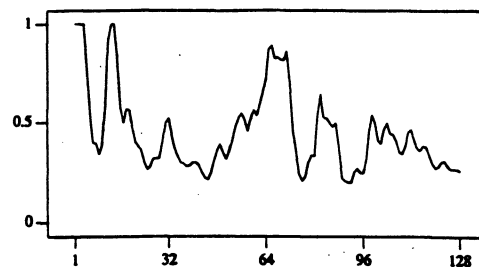


Fig. 4. A scanline taken from a digitized natural image. There are 128 pixels in the scanline.

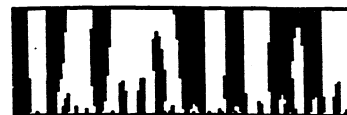


Fig. 5. The sign of the scale-space filtered data shown in Fig. 4. Positive regions are shown in white, negative in black, and values very near zero are in gray. The zero crossings occur on the borders between black and white regions, and along contours within the gray regions.

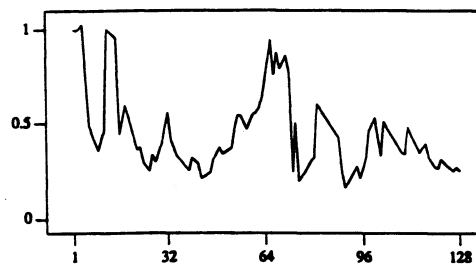


Fig. 6. A reconstruction f of a signal from the reconstructed Laplacian-of-Gaussian filtered data v . This function should be the same as Fig. 4 if stable reconstruction from zero crossings were possible.

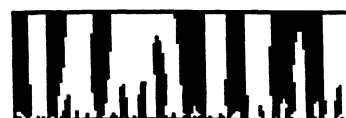


Fig. 7. The sign of the scale-space filtered data shown in Fig. 4. The data are nearly identical to those of Fig. 5, showing that the signals in Figs. 4 and 6 have virtually the same zero crossings for their Laplacian-of-Gaussian filtered data at all scales of resolution.

should be noted that the sign violations are small: the rms average value of the violations is only 10^{-4} of the rms average value of v .) The sum of levels was used to obtain the reconstructed function, by the pseudoinverse method described in Section IV-C. This reconstructed function is shown in Fig. 6; and Fig. 7 shows its sign of Laplacian monolith.

It is evident that the zero crossings of the original and reconstructed functions are essentially identical. Also, major features of the signal are reproduced, although there are significant differences in the small-scale details. We conclude that while the reconstruction from zero crossings preserves some of the essential information about the signal, it is unstable with respect to high-frequency components.

Next, the reconstruction process was applied to the same data, but this time holding fixed the values of σ along the zero crossings of v_0 , and leaving free the values of v

at maximum blur. In this way, we have specified the zero crossings and the gradient information at the zero crossings through 40 levels in scale space. After approximately 6000 iterations, the equation error plus penalty fell below 10^{-16} . The reconstructed function is not shown because it is visually indistinguishable from the original signal. The reconstruction function differed from the original by less than 10^{-6} at all points (on an intensity scale of 0 to 1). This result can be taken as empirical evidence that the representation based on zero crossings plus gradient data for one-dimensional images is complete and stable.

B. Two-Dimensional Experiments

We next turn our attention to two-dimensional images. The original images are 64×64 pixels with 8-bit intensity resolution. Because of computing constraints, only 20 levels of scale were used in the reconstructions. Each successive level is obtained from the previous by means of the 2-D discrete Heat Equation, given in Section IV-B-2. Fig. 8 shows an original image of a pagoda and the result of reconstruction from zero crossings. The equation-error minimization was run for approximately 2000 iterations. The reconstruction was computed by subtracting the sum of levels of the reconstructed Laplacian monolith from the blurred original image, as described in Section IV-C. As can be seen, the two images are perceptually different. Mainly, the high-frequency content is not well reconstructed, so that the reconstruction appears as a slightly blurred version of the original. Despite the evident differences between the two images, however, the zero crossings of the two images are nearly identical. There are sign differences at less than 2 percent of the points in the 20-layer monolith, concentrated in the bottom (high-frequency) layers. Virtually all of these sign differences are adjacent to zero crossings, causing at most a one-pixel shift in the location of the zero crossing.

The reconstruction by equation error minimization is sensitive to the initial guess used to start the conjugate-gradient iteration. For the reconstruction in Fig. 8(b), the initial Laplacian monolith contained values of ± 1 everywhere, with the sign matching the required sign at each point. The input image magnitude was in the range $[0, 255]$. Fig. 9 shows the results obtained using initial values of ± 256 instead. Even after running for nearly 20 000 iterations, this reconstruction shows considerable noise and artifacts. These artifacts take the form of strong positive values side-by-side with strong negative values, which largely cancel each other out in the blurring process. Once again, the zero crossings of the reconstructed image in Fig. 9 are nearly the same as those of the original image. The differences between the reconstructed images (Figs. 8(b) and 9) and the original [Fig. 8(a)] can be taken as an empirical demonstration of the lack of stability of the representation based on zero crossings alone.

To show the similarities of the zero-crossing representations, we show the sign of the Laplacian-of-Gaussian for the original image and the reconstructed images, in Fig. 10, at a sampling of levels. The zero crossings are

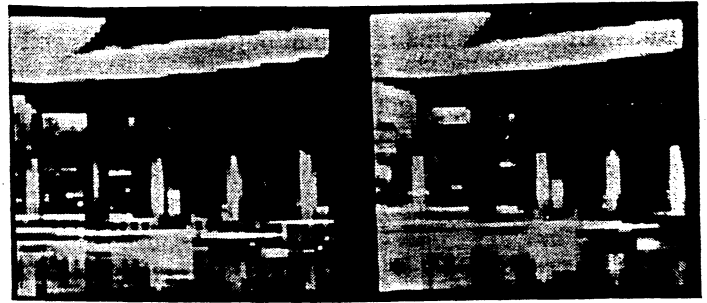


Fig. 8. An original image, of a Japanese pagoda, and a reconstruction from the zero crossings in scale space.



Fig. 9. Another reconstruction of the pagoda image of Fig. 8(a), using a different initialization.

consequently at the borders between black and white regions. As can be seen, there are some differences between the zero crossings, especially at the bottom levels. However, as noted before, the differences consist almost exclusively of a one-pixel displacement in the location of the zero crossings. This is more easily seen when one displays the zero crossings as curves, and flickers between the two images.

It is worth emphasizing that the sign of Laplacian-of-Gaussian in Fig. 10(b) was computed from the reconstructed image, and is not the sign of the reconstructed monolith. The solution method involves the reconstruction of a function in scale space, \tilde{v} , which approximately satisfies the heat equation, subject to a penalty term which tends to make $\text{sgn}(\tilde{v})$ agree with the given $\text{sgn}(v_0)$. The resulting \tilde{v} , summed over scale space, is used to reconstruct an approximation \tilde{f} to the original image f . The Laplacian-of-Gaussian computed from \tilde{f} , call it $v_{\tilde{f}}$, should approximate v_0 , the Laplacian-of-Gaussian of the original image f . However, each of the scale-space functions v_0 , \tilde{v} , and $v_{\tilde{f}}$, is different. It is the sign of $v_{\tilde{f}}$ which is shown in Fig. 10, representing the true locations of the zero crossings of \tilde{f} . Accordingly, Figs. 8(a) and (b) and 9 represent three images whose zero crossings are nearly the same at all levels, and yet have substantially different numerical and perceptual values. Thus, similarly to the one-dimensional case described earlier, these results empirically establish the instability of the zero-crossing representation, when zero crossings are used alone. Of course, one should modify this broad statement with the note that a particular method of handling borders has been specified, and the image comes from a particular class of images.

The image in Fig. 8(a) contains a large number of sharp edges, and thus high-frequency components. It is known

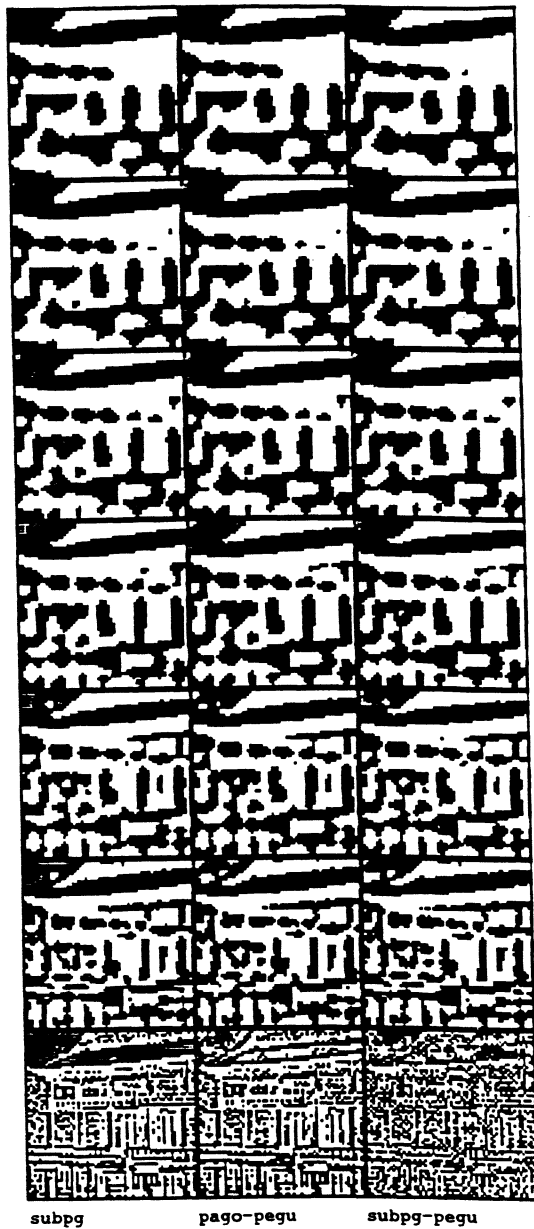


Fig. 10. Levels 0, 3, 6, 9, 12, 15, and 18 of the sign of the Laplacian-of-Gaussian "monolith" of (left column) the original image, (middle column) the reconstructed image Fig. 8(b), and (right column) the reconstructed image Fig. 9.

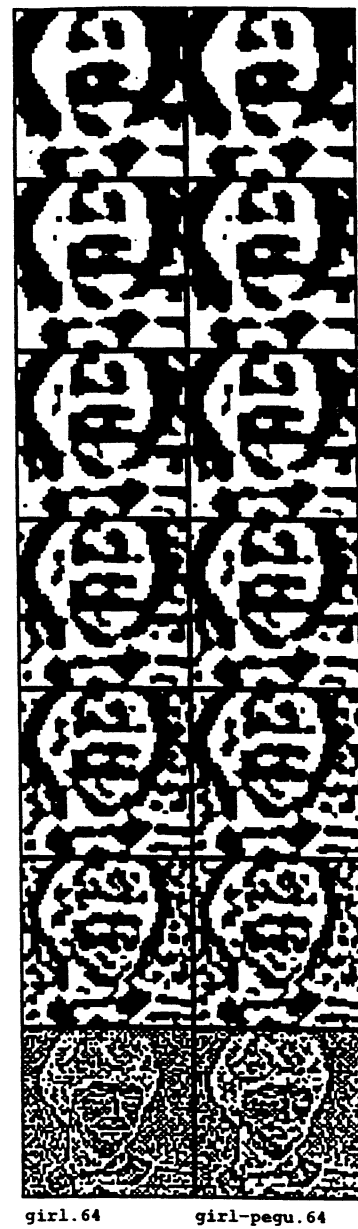


Fig. 12. Levels 0, 3, 6, 9, 12, 15, and 18 of the sign of the Laplacian monolith of (left column) the original image, (second column) the reconstructed image of Fig. 11.

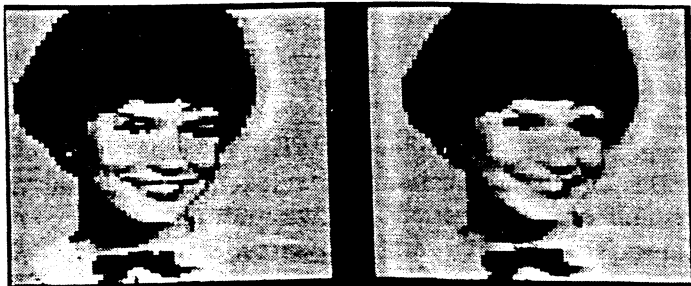


Fig. 11. An original image, and the reconstruction using the zero crossings. The original image is relatively free of high-frequency components.

that in the conjugate-gradient method of minimizing equation error, the low-frequency components will converge before the high frequencies. In Fig. 11, we show a standard test image. The image has significantly smaller

Fourier components (nearly a factor of 10, on average) at the higher frequencies as compared to the image of the pagoda, in Fig. 8. The reconstruction from zero crossings of this image works quite well, as shown in Fig. 11(b). Again there are in fact some significant numerical differences between the original and reconstruction in Fig. 11, but the reconstructed image is perceptually similar to the original, except for fine details. The signs of the Laplacian-of-Gaussians of the images are shown in Fig. 12, and are essentially the same, as in Fig. 10. Once again, there are minor differences at the low levels (high frequencies).

In another experiment, not shown, using an even softer original image in which fine detail was relatively lacking, the reconstructed image was perceptually almost identical to the original. Further, the iterative process for reconstructing I' converged somewhat more quickly to a low



Fig. 13. A reconstruction of the pagoda image [Fig. 8(a)] using zero crossings and gradient data along the zero crossings.



Fig. 14. A reconstruction of the pagoda image [Fig. 8(a)] using zero crossings and gradient data along zero crossings, with the gradient contaminated by noise of relative magnitude $1/256$ as described in the text.



Fig. 15. A reconstruction of the pagoda image [Fig. 8(a)] using zero crossings and approximately 10 percent of the gradient data along the zero crossings.



Fig. 16. A reconstruction of the pagoda image [Fig. 8(a)] using zero crossings together with data along zero crossings in the bottom two levels only.

equation error with this image than with the image of the pagoda.

In Fig. 13, we show the result of a reconstruction of Fig. 8(a), where the representation by zero crossings has been enhanced with the value of the gradients along the zero crossing. The reconstruction is now stable, accurate, and fast. There are minor numerical differences, but no major perceptual differences between the original and the reconstruction. Of course, the amount of data in the entire representation exceeds the information content in the original image, since at every zero-crossing point on every level, a value of the magnitude of the gradient of the filtered image must be stored. Since these values are stored as floating-point numbers, the entire data structure is in no way a compaction of the original image. However, a compact code is not the objective of the representation by zero crossings, so that at this junction, we can conclude that a representation that includes, with the zero crossings, the gradient data along the zero crossings is empirically complete and stable.

In other experiments using gradient data, we have found that reconstruction is stable for all kinds of images. Noise can be added to the gradient data, and stable reconstruction is still not difficult. For Fig. 14, multiplicative noise in the range of $1 \pm 1/256$ was applied to the gradient values along the zero crossings before reconstruction. The noise simulates storing the gradient values with 8-bit precision, and still yields a good representation. However, even with 8-bit values for the gradients, the total amount of data in the typical representation exceeds the data content of the original image.

In Fig. 15, we show the same image reconstructed from a representation using the zero crossings together with 10 percent of the gradient values. The locations of the gra-

dent values were chosen at random from the set of all zero-crossing pixels. Reconstruction is much improved from the zero crossings alone. However, using only 10 percent of the gradient values, some noise is present in the reconstruction. If the percentage of gradient values retained is increased, then the reconstruction is improved. Indeed, with 50 percent of the gradient values, reconstruction is approximately as good as in Fig. 13, where 100 percent of the gradient values were retained. Good reconstructions can also be obtained by concentrating the gradient data in the higher-frequency levels. Fig. 16 shows a reconstruction for which gradients were specified on all zero crossings in the bottom two levels, and nowhere else. The number of zero crossings at which gradients were specified came to about 22 percent of the total. The reconstruction is nearly as good as that for which gradients were completely specified at all levels. This result shows how the instability is greatest for the high frequencies: when these are suitably controlled, the reconstruction proceeds stably.

VI. CONCLUSIONS

Despite the enthusiasm for zero crossings in multi-resolution representations that was prevalent in the computer vision research community in the early 1980's, their use has declined considerably, in favor of other edge detection and feature detection techniques. Where zero crossings are still used, the system generally incorporates some measure of the gradient along the zero crossings, such as a thresholding operation to eliminate zero crossings with low gradients.

The analyses and results in this paper tend to corroborate the wisdom of this trend. Specifically, we have seen that there exist distinct images with quite similar zero-

crossing surfaces (at least through a large range of scale space), giving an empirical example of the instability of zero crossings alone as an intermediate representation. We have also demonstrated, again empirically, that reconstruction from zero crossings together with gradient values along the zero crossings is possible and stable. The chief disadvantage of this representation is that there is much redundancy in the bulk of data included in the representation.

The many mathematical results showing that the zero crossings form a complete representation under analytical assumptions, we have seen, are primarily based on analytic continuation results, and thus suggest an unstable representation. The likely situation is that the zero crossings alone are enough to be complete, in most cases, but insufficient to provide stability. One might hope that the inclusion of a little additional information will make the entire representation stable, but short of adding all the gradient information along the zero crossings, this remains to be seen. Preliminary experiments are not encouraging.

The mathematical analysis of a representation frequently will occur after the technique has been proven in-the-field, or has been largely discarded. Such is probably the case with the analysis of zero crossings for computer vision applications. We could hope to establish a mathematical analysis of the stability of these representations, rather than merely the empirical results cited here, but for vision applications these theorems would be largely irrelevant.

The technique of minimizing equation error for solving the inverse problem of reconstruction from a representation, however, seems to yield far better results than methods involving the minimization of data error, at least for the applications studied here. We expect that the technique will be useful for other inverse problems, and presumably some in the computer vision field. The general methodology of studying a representation in terms of its mathematical properties, and developing reconstruction methods to evaluate the stability and variations in the fibers, in analogy with the study undertaken here, is highly recommended.

APPENDIX

We show that if two irreducible polynomials with real coefficients in n real variables share in common the same ($n - 1$ -dimensional) zero crossing (even for just an open patch), then the two polynomials are the same. This proof was supplied by Mumford. We first prove an analog of the decomposition theorem for real algebraic varieties.

Proposition 1: Let $\pi \subset \mathbb{R}[x_1, x_2, \dots, x_n]$ be a prime ideal and let $h =$ the height of π (i.e., the length of a maximal chain of prime ideals:

$$(0) = \pi_0 \subsetneq \pi_1 \subsetneq \pi_2 \subsetneq \dots \subsetneq \pi_h = \pi.$$

Then if $V(\pi)$ is the locus of zeros in \mathbb{R}^n shared by all polynomials in the ideal π (the real variety of the ideal π), then $V(\pi)$ is the union of real submanifolds of \mathbb{R}^n of dimension $n - h$ and less.

Proof: By the standard decomposition theorem, the complex variety $V(\pi) \subseteq C^n$, which is the locus of complex zeros of the polynomials in π viewed as polynomials of n complex variables, breaks up into a smooth part, which is always an $n - h$ -dimensional complex submanifold of C^n , and a singular part, which is a finite union of subsets of the form $V(\pi')$, with π' being a prime ideal satisfying $\pi \subsetneq \pi'$ (and thus π' is of greater height). The smooth part intersects \mathbb{R}^n in an $n - h$ -dimensional real submanifold of \mathbb{R}^n , and so the proposition follows by decreasing induction on the height of the prime ideal. The main result is Proposition 2.

Proposition 2: Let $p \in \mathbb{R}[x_1, x_2, \dots, x_n]$ be an irreducible polynomial. Thus, the ideal $\pi = (p)$ generated by p has height one. Let $x \in V(p)$ be a smooth point (i.e., a point where $V(p)$ is an $n - 1$ -dimensional submanifold), and let U be an open neighborhood of x in \mathbb{R}^n . Then for any real polynomial q of n variables, if

$$q = 0 \text{ on } U \cap V(\pi),$$

then p divides q .

Proof: Let I be the ideal of all polynomials in $\mathbb{R}[x_1, x_2, \dots, x_n]$ that vanish on $U \cap V(\pi)$, i.e., the ideal defined by this subset. The radical of I , denoted \sqrt{I} , consists of all polynomials such that some power of the polynomial lies in I . A standard theorem in algebraic geometry, used in the proof of a strong form of the Nullstellensatz, shows that the radical of I is the finite intersection of prime ideals containing I , so that

$$\sqrt{I} = \pi_1 \cap \pi_2 \cap \dots \cap \pi_k, \pi_i \text{ prime ideals.}$$

Now, the patch $U \cap V(\pi)$ is contained in $V(\sqrt{I})$, by the definition of I . So,

$$U \cap V(\pi) \subset V(\sqrt{I}) = V(\pi_1) \cup V(\pi_2) \cup \dots \cup V(\pi_k).$$

Further, $p \in I \subset \sqrt{I}$, so that $p \in \pi_i$ for all i . Suppose that $(p) \subsetneq \pi_i$ for every i . Then the height of each π_i is at least two, and thus by proposition 1, the variety of each $V(\pi_i)$ consists of submanifolds of dimension $n - 2$ and less. Thus, the union of the $V(\pi_i)$'s cannot contain an $n - 1$ -dimensional patch, which is a contradiction. So for some i , we must have $(p) = \pi_i$, and so

$$(p) \subset I \subset \pi_1 \cap \dots \cap \pi_k \subset \pi_i = (p).$$

Recalling that q is in I , we have therefore that $q \in (p)$, i.e., that p divides q , as required. ■

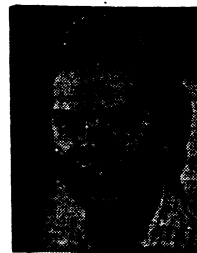
As a result of the proposition, we have a uniqueness theorem. Suppose that p and q are two polynomials with domain \mathbb{R}^n . Suppose furthermore that Δp and Δq are irreducible polynomials. The corresponding scale-space functions v_1 and v_2 will also be polynomials (of $n + 1$ variables) and must be irreducible, for otherwise a factorization of either would induce (at $t = 0$) a factorization of the Laplacian of the initial polynomial. Suppose that v_1 and v_2 share in common an open patch of some zero crossing. Recall that the zero crossings are defined in terms of the boundaries of the regions of negativity and

positivity, and thus will be submanifolds of codimension one. Applying the proposition leads to the conclusion that v_1 and v_2 divide one another, or equivalently, are equal. Consequently, p and q are equal.

In fact, all that is necessary for uniqueness is that the scale-space functions v_1 and v_2 should be irreducible. To extend the results, it would be interesting to characterize all factorable polynomials that satisfy the Heat Equation.

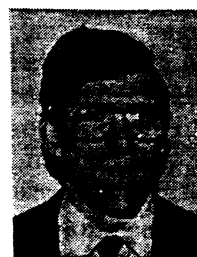
REFERENCES

- [1] D. Marr, *Vision*. San Francisco, CA: W. H. Freeman, 1982.
- [2] H. Barrow and J. M. Tenenbaum, "Recovering intrinsic scene characteristics from images," in *Computer Vision Systems*, A. Hanson and E. Riseman, Eds. New York: Academic, 1978.
- [3] J. Richter and S. Ullman, "Non-linearities in cortical simple cells and the possible detection of zero crossings," *Biological Cybern.*, vol. 53, pp. 195-202, 1986.
- [4] A. Rosenfeld and M. Thurston, "Edge detection and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. C-20, pp. 562-569, 1971.
- [5] P. Burt and T. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, p. 532, 1983.
- [6] A. Witkin, "Scale space filtering," in *Proc. 8th Int. Joint Conf. Artificial Intell.*, 1983, p. 1019.
- [7] J. J. Koenderink, "The structure of images," *Biological Cybern.*, vol. 50, pp. 363-370, 1984.
- [8] S. W. Zucker and R. A. Hummel, "Receptive fields and the representation of visual information," in *Proc. 7th Int. Conf. Pattern Recogn.*, July 1984.
- [9] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 26-33, 1986.
- [10] A. L. Yuille and T. A. Poggio, "Scaling theorems for zero crossings," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 15-25, 1986.
- [11] L. Kovaszny and H. Joseph, "Image processing," *Proc. IRE*, vol. 43, pp. 560-570, 1955.
- [12] C. Enroth-Cugell and J. Robson, "The contrast sensitivity of retinal ganglion cells of the cat," *J. Physiol.*, p. 517, 1966.
- [13] S. Laughlin, "Matching coding to scenes to enhance efficiency," in *Physical and Biological Processing of Images*, O. Braddick and A. Sleigh, Eds. New York: Springer, 1983.
- [14] R. Hummel, "The scale-space formulation of pyramid data structures," in *Parallel Computer Vision*, L. Uhr, Ed. New York: Academic, 1987, pp. 107-123.
- [15] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. London (B)*, p. 187, 1980 (see appendix).
- [16] J. G. Daugman, "Six formal properties of anisotropic visual fibers," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, pp. 882-887, 1983.
- [17] K. Prazdny, "Stereopsis in the absence of zero-crossings in the band-pass filtered images," Memo, Lab. AI Res., Fairchild, Schlumberger, 4001 Miranda Ave., Palo Alto, CA 94304.
- [18] D. Marr and T. Poggio, "A computational theory of human stereo vision," *Proc. Roy. Soc. London (B)*, p. 301, 1979.
- [19] B. Kedem, "Zero-crossings analysis," Tech. Rep. TR85-38, Dep. Math., Univ. Maryland, College Park, MD 20742, Nov. 1985.
- [20] P. Johansen, S. Skelboe, K. Grue, and J. D. Anderson, "Representing signals by their topoints in scale space," in *Proc. 8th Int. Conf. Pattern Recogn.*, vol. 1, 1986, pp. 215-217.
- [21] L. Bers, F. John, and M. Schechter, *Partial Differential Equations*. Providence, RI: American Mathematical Society, 1964.
- [22] M. Protter and H. Weinberger, *Maximum Principles in Differential Equations*. Englewood Cliffs, NJ: Prentice-Hall, 1967.
- [23] F. John, *Partial Differential Equations*. New York: Springer-Verlag, 1975.
- [24] R. A. Hummel and B. C. Gidas, "Zero crossings and the heat equation," NYU Robotics Rep. 18, Mar. 1984.
- [25] A. L. Yuille and T. Poggio, "Fingerprints theorems," *Proc. Amer. Assoc. Artificial Intell.*, pp. 362-365, 1984.
- [26] W. Rudin, *Functional Analysis*. New York: McGraw-Hill, 1972.
- [27] R. A. Hummel, B. Kimia, and S. Zucker, "Deblurring Gaussian blur," *Comput. Vis., Graphics, Image Processing*, vol. 38, pp. 66-80, 1987.
- [28] B. Logan, "Information in the zero-crossings of band pass signals," *Bell Syst. Tech. J.*, vol. 56, p. 510, 1977.
- [29] J. Sanz and T. Huang, "Image representation by sign information," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 729-738, 1989.
- [30] S. Curtis, A. Oppenheim, and J. Lim, "Signal reconstruction from Fourier transform sign information," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 643-657, 1985.
- [31] A. L. Yuille and T. Poggio, "Fingerprints theorems for zero crossings," *J. Opt. Soc. Amer.*, vol. 2, pp. 683-692, 1985.
- [32] S. Curtis, S. Shitz, and A. Oppenheim, "Reconstruction of nonperiodic two-dimensional signals from zero crossings," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 890-893, 1987.
- [33] S. Curtis and A. Oppenheim, "Reconstruction of multidimensional signals from zero crossings," *J. Opt. Soc. Amer.*, vol. 4, pp. 221-230, 1987.
- [34] Y. Y. Zeevi and D. Rotem, "Image reconstruction from zero crossings," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1269-1277, 1986.
- [35] R. Hummel and R. Moniot, "Solving ill-conditional problems by minimizing equation error," in *Proc. IEEE 1st Int. Conf. Comput. Vision*, 1987, pp. 527-533.
- [36] —, "A network approach to reconstructions from zero-crossings," in *Proc. IEEE Comput. Soc. Workshop Comput. Vision*, Nov. 1987, pp. 8-13.
- [37] B. Lowe, "A variational method for parameter identification," Ph.D. dissertation, New York Univ., Oct. 1986.
- [38] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*. New York: Academic, 1980.
- [39] D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 413-424, 1986.
- [40] M. R. Hestenes, *Conjugate Direction Methods in Optimization*. New York: Springer-Verlag, 1980.
- [41] G. Strang and G. Fix, *An Analysis of the Finite Element Method*. Englewood Cliffs, NJ: Prentice-Hall, 1973.



Robert Hummel (M'82) received the B.A. degree in mathematics from the University of Chicago, and the Ph.D. degree in mathematics from the University of Minnesota.

He has studied at the Computer Vision Lab of the Computer Science Center at the University of Maryland, doing research on computer vision with Professor A. Rosenfeld. While there, he became one of the co-developers of relaxation labeling methods for scene analysis. He has subsequently worked at Stanford Research Institute, Honeywell's System and Research Center in Minneapolis, MN, Martin Marietta's Signal and Image Processing Lab in Orlando, and as an Adjunct Lecturer at IBM's System Research Institute in New York. His work in image processing includes topics in image enhancement, operator and filter design, autonomous target recognition, and applications of relaxation labeling. Since 1980 he has been on the Faculty of New York University's Courant Institute of Mathematical Sciences, first as a postdoc in Mathematics, and then on the Faculty of the Department of Computer Science, where he is currently an Associate Professor. He is a member of the Courant Institute's Laboratory for Robotics Research and Computer Vision, and is also a member of New York University's Washington Square Center for Neural Science. During 1989 he is a Visiting Professor at Vrije University in Amsterdam, The Netherlands.



Robert Moniot received the B.S. degree in physics from the State University of New York, College at Fredonia, the Ph.D. degree in physics from the University of California, Berkeley, and the M.S. degree in computer science from New York University.

In 1982, following postdoctoral work at the Nuclear Physics Laboratory of Rutgers University, he took up his present position on the Faculty of the Division of Science and Mathematics, Fordham University, College at Lincoln Center. He has worked with R. Hummel, at the Robotics Laboratory of the Courant Institute, on the application of numerical methods to image enhancement and image reconstruction.