Nonlinear Image Estimation Using Piecewise and Local Image Models

Scott T. Acton, *Member, IEEE,* and Alan C. Bovik, *Fellow, IEEE*

***Abstract*—We introduce a new approach to image estimation based on a flexible constraint framework that encapsulates meaningful structural image assumptions. Piecewise image models (PIM's) and local image models (LIM's) are defined and utilized to estimate noise-corrupted images. PIM's and LIM's are defined by image sets obeying certain piecewise or local image properties, such as piecewise linearity, or local monotonicity. By optimizing local image characteristics imposed by the models, image estimates are produced with respect to the characteristic sets defined by the models. Thus, we propose a new general formulation for nonlinear set-theoretic image estimation. Detailed image estimation algorithms and examples are given using two PIM's: piecewise constant (PICO) and piecewise linear (PILI) models, and two LIM's: locally monotonic (LOMO) and locally convex/concave (LOCO) models. These models define properties that hold over local image neighborhoods, and the corresponding image estimates may be inexpensively computed by iterative optimization algorithms. Forcing the model constraints to hold at every image coordinate of the solution defines a *nonlinear regression* problem that is generally nonconvex and combinatorial. However, approximate solutions may be computed in reasonable time using the novel generalized deterministic annealing (GDA) optimization technique, which is particularly well suited for locally constrained problems of this type. Results are given for corrupted imagery with signal-to-noise ratio (SNR) as low as 2 dB, demonstrating high quality image estimation as measured by local feature integrity, and improvement in SNR.**

***Index Terms*—Image enhancement, image estimation.**

I. Introduction

ONE OF THE oldest ongoing problems in image processing is *image estimation*, which encompasses algorithms that attempt to recover images (usually digital) from observations. More specifically, it is generally desired to remove unwanted noise artifacts, which are often broadband, while simultaneously retaining significant high-frequency image features, such as edges, texture and detail. In such a context, the problem is often referred to as *image enhancement.* The objectives of image estimation/enhancement are generally twofold, and conflicting: smoothing of image regions where the intensities vary slowly, and simultaneous preservation of sharply-varying, meaningful image structures.

The first main theme of the current paper is the development of image estimation algorithms that *begin with a model for the image.* The model used should, of course, be designed to capture meaningful image detail and structure for the application at hand. We explore several fairly general image models that are based on well-defined local image characteristics. The models that we study are divided into two classes: *piecewise image models* (PIM's), which model images as everywhere obeying a certain property (such as constancy or linearity) in a piecewise manner, and *local image models* (LIM's), which characterize images as obeying a certain property (such as monotonicity or convexity) over every subimage of specified geometry.

A second main theme of the paper is the casting of the estimation problem as an *approximation* to a *nonlinear regression* with respect to the *characteristic set* defining the image model. Estimation proceeds by encouraging adherence to the model properties while maintaining a semblance (a minimum distance) to the observed input image. The goal is to compute a solution image that approximates the desired image property and that also is at minimum distance (defined by a prescribed distance norm) from the observed image.

The approach to image estimation described here is generally quite new. Some related methods have been reported that attempt to preserve image smoothness in a more usual sense (small derivative or Sobolev image norm), while at the same time producing an output image that is "close" to the input image [3], [10], [11], [13]. In these constrained optimization or regularized methods, the smoothness constraint can be relaxed at image boundaries—identified via *line processes* [10]. The regions between the discontinuities can be modeled as weakly continuous surfaces, using a weak membrane model [4] or a two-dimensional (2-D) noncausal Gaussian Markov random field (GMRF) model [12], [23]. These approaches, while often effective, do suffer from some drawbacks. First, they do not fall within a flexible, unified framework that allows for the use of different image models demanded by different applications. Second, the implementation of smoothness constraints that decouple across intensity boundaries is somewhat difficult (since the estimation of line processes is a hard problem), whereas models such as local monotonicity and piecewise linearity naturally preserve boundaries between smooth regions. Finally, the computational cost of obtaining image estimation results using constrained combinatorial optimization is impractical for time-critical image processing applications. Here it is shown that approximate nonlinear regression with respect

Manuscript received November 30, 1995; revised September 10, 1997. This material is based on work supported in part by the U.S. Army Research Office under Grant DAAH04-95-1-0255. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Moncef Gabbouj.

S. T. Acton is with the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: sacton@master.ceat.okstate.edu).

A. C. Bovik is with the Laboratory for Vision Systems, Center for Vision and Image Sciences, Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712-1084 USA (e-mail: bovik@ece.utexas.edu).

Publisher Item Identifier S 1057-7149(98)04371-1.

to PIM's and LIM's can be accomplished with relatively low computational complexity via the recently introduced generalized deterministic annealing (GDA) algorithm. GDA is a starting-state independent iterative optimization technique that is particularly well-suited for *locally constrained* problems such as those studied here.

The paper is organized as follows. Section II outlines the nonlinear regression approach to nonlinear image estimation. Sections III and IV describe four image models (two PIM's and two LIM's) and the estimation procedure in each case. Computed image estimation examples are provided for each model. Section V briefly discusses the iterative optimization algorithm GDA, particularly those aspects that relate to the set-theoretic image estimation problem. The utilization of GDA leads to a nonheuristic implementation that is particularly efficient for the problem. The paper is concluded in Section VI.

II. Nonlinear Image Estimation

A. Nonlinear Image Estimation and the Relationship to Nonlinear Regression

Consider the problem of estimating a discrete-space image $\mathbf{i} = [i(x, y); 0 \leq x, y \leq N-1, 0 \leq i(x, y) \leq K-1, i(x, y) \in \mathbf{Z}]$ from an observed image $\mathbf{g}$ where

$$\mathbf{g} = \mathbf{i} + \mathbf{n}. \tag{1}$$

In (1), $\mathbf{n}$ represents additive independent, identically distributed (i.i.d.) noise where $\mathbf{n} = [n(x, y); 0 \leq x, y \leq N-1, n(x, y) \in \mathbf{Z}]$.

The image estimation problem posed by (1) can be solved via *nonlinear regression*

$$\mathbf{g}^* = \arg \min_{\mathbf{h} \in \mathbf{C}} ||\mathbf{g} - \mathbf{h}||_D. \tag{2}$$

Here, the optimizing estimate $\mathbf{g}^*$ is the (generally nonunique [18]) image closest to the observation $\mathbf{g}$, among all images that lie *within* the *characteristic set* $\mathbf{C}$, i.e., images that strictly satisfy the image model (a PIM or LIM, in this case). The characteristic set $\mathbf{C}$ defines the *characteristic property* of the regression, such as local monotonicity or piecewise constancy. The term $||\mathbf{g} - \mathbf{h}||_D$ is the distance between image $\mathbf{h}$ and the observed image $\mathbf{g}$, defined by an appropriate distance norm $|| \cdot ||_D$.

Solving (2) is generally an expensive combinatorial optimization for data sets approaching the size of images [18], [19]. Locally monotonic regression algorithms in [18] are of exponential complexity, although a recent algorithm that promises linear complexity when operating on signals from a finite alphabet has been proposed [22]. In the current paper, we take a different approach: we recast the problem by treating membership in $\mathbf{C}$ as a soft constraint. This leads to a problem well-suited to fast optimization algorithms.

B. Existence and Statistical Optimality of Nonlinear Regression

Nonlinear regression of the form (2) always has at least one solution provided that the characteristic set $\mathbf{C}$ is a closed set [18], as in all the cases considered here. Nonlinear regression also has an interpretation as projection of the signal to be regressed onto the characteristic set $\mathbf{C}$. The projection is with respect to a semimetric [18]. The geometrical structure of the regression problem also admits a strong statistical optimality property. Indeed, if the additive noise $\mathbf{n}$ in (1) consists of i.i.d. samples coming from a discrete version of the generalized exponential distribution function $[p \in (0, \infty)]$ with density

$$f_p(x) = \gamma \cdot \exp(-\zeta |x|^p) \tag{3}$$

then the solution to the nonlinear regression (2) is a maximum likelihood estimate, provided that the distance norm used is the p-semimetric [20]

$$||\mathbf{g} - \mathbf{h}||_D = \left[\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |g(x, y) - h(x, y)|^p \right]^{1/p}. \tag{4}$$

Thus, if the image noise can be modeled as i.i.d. and coming from the density (3) for some p, then the nonlinear regression problem can be formulated as maximum likelihood via selection of the norm (4).

The generalized exponential distribution includes three very common additive noise models that will be employed here. For $p = 1$, $f_p(x)$ is the Laplacian density, and the optimizing data constraint leading to an ML estimate is the l^1-norm. Laplacian noise is a common "heavy-tailed" or highly impulsive noise model, e.g., to model data containing outliers. For $p = 2$, $f_p(x)$ is Gaussian density, and the ML estimate is under the l^2-norm. Finally, as $p \to \infty$, $f_p(x)$ becomes uniform density, and the distance measure to use is the l^∞-norm.

We can use the preceding observations to guide the selection of the norm $||\mathbf{g} - \mathbf{h}||_D$ in the construction of a cost (energy) functional for a regularized solution. The regularized solution encapsulates soft constraints for consistency with the sensed image $\mathbf{g}$ and adherence to the characteristic property. Although the introduction of the soft model constraint to replace the hard constraint that the solution lie in $\mathbf{C}$ changes the problem, if the solution is forced toward both the original image under the appropriate data constraint norm and also toward the characteristic set, an estimate $\hat{\mathbf{g}}$ of the optimal regression will be obtained which may be more physically sensible than the regression $\mathbf{g}^*$.

C. Regularized Solution

In the regularized solution, the image estimate $\hat{\mathbf{g}}$ is found by minimizing an energy functional $E(\mathbf{h})$ that combines a penalty for deviation from the observed image data with a penalty for local deviation from the characteristic image property—assumed to be a PIM or a LIM:

$$E(\mathbf{h}) = ||\mathbf{g} - \mathbf{h}||_D + \lambda ||\mathbf{h} - \mathbf{PROP}(\mathbf{h})||_M. \tag{5}$$

Thus

$$\hat{\mathbf{g}} = \arg \min_{\mathbf{h}} \{E(\mathbf{h})\}. \tag{6}$$

In (5), $||\mathbf{g} - \mathbf{h}||_D$ is the distance between image $\mathbf{h}$ and the observed image $\mathbf{g}$, as defined in (4). This term is called the *data constraint.* The distance norm is generally motivated by

a priori information about the noise process $\mathbf{n}$, as described in Section II-B. In all of the simulations, additive noise from the general density (3) will be used for $p \in \{1, 2, \infty\}$. In each case, the appropriate optimal p-norm or p-semimetric is used to define the data constraint.

The term $||\mathbf{h} - \mathbf{PROP}(\mathbf{h})||_M$ in (5), the *model constraint*, provides an energy penalty for local deviation from a *characteristic property* $\mathbf{PROP}$ which defines the image model. The form of the *model norm* $\|\cdot\|_M$ depends on the characteristic property. However, in general it will be written

$$||\mathbf{h} - \mathbf{PROP}(\mathbf{h})||_M = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |h(x, y) - PROP[h(x, y)]| \tag{7}$$

where $|h(x, y) - PROP[h(x, y)]|$ is a local measure of error energy relative to the characteristic set.

The characteristic properties studied here will be defined by PIM's and LIM's. The model constraint is computed by summing, over all image coordinates, the absolute distance between $\mathbf{h}$ and the closest local solution to $\mathbf{h}$ that satisfies the characteristic property locally. Again, a suitable distance norm may be selected to define the model constraint according to some statistical or structural criterion.

The regularized solution (6) is a more tractable approximation to the regression (2). However, aside from the issue of computational complexity, it can be argued that (6) may often present a more physically sensible solution than (2). Consider the case of (5), where λ is taken to be large: the model constraint is thus given considerably greater weight than data constraint. If λ is taken sufficiently large, then the solution image will be forced to adhere to the characteristic property at nearly every, and possibly at all image coordinates. In such a case, the solution may not adequately resemble the input image in some locations, owing to local deviations from the image model. It may therefore be argued that the nonlinear regression (2) yields solutions which may be numerically optimal, yet suboptimal in the important sense of *image enhancement.*

The *regularization parameter* λ determines the degree to which $\hat{\mathbf{g}}$ will conform to the data constraint versus the model constraint. In [8], methods were explored for determining such relative weights for more usual (linear) smoothness operators. Generally, the estimation of λ depends on the *a priori* knowledge of the corruptive noise and is typically complex and time-consuming. Because operators used to evaluate the characteristic properties (the PIM's and LIM's) are nonlinear (unlike the traditionally used Laplacian operator), the methods used in [8] are not applicable to this implementation. Instead, the regularization parameter may be selected via cross validation [17]. With this method, the image is first divided into an estimation set and a validation set. To evaluate the solution quality given for a particular regularization parameter, the nonlinear image estimation is performed using (5) on the pixels in the estimation set. Simultaneously, image estimation is implemented for the pixels in the validation set, but with a cost functional that does not include the data constraint. So, the pixels in the validation set can be used to predict the estimation error [17]. The main drawback of using cross validation to select the regularization parameter is that the cost to evaluate a particular λ is equivalent to the cost of performing image estimation itself.

Empirically, we have found the image estimation procedure to be quite robust with respect to selection of λ; indeed, values of λ that differ by one or two orders of magnitude (10 or 100) do not yield very different results than obtained here. This is due to the fact that the constraints defined by the PIM's and LIM's used here are fully realizable. Meaningful image estimates can be computed that have zero cost penalties from the PIM and LIM constraints, in contrast to the Laplacian operator that produces a zero-energy penalty only for an image without edges. The results demonstrate this—in every example given in the paper, over 90% of the pixels in the obtained image estimate obey the defining characteristic property. However, algorithms of this type appear to be somewhat sensitive to under-specification of λ—for values an order of magnitude smaller than unity (thus heavily weighting the data constraint relative to the model constraint), the solution quality begins to deteriorate.

Note that in the absence of *a priori* information concerning the original image structure, cross validation may be also applied to select the appropriate PIM or LIM for image estimation. With this approach, the validation error [17] (the predicted mean-squared error) is computed for each potential model using the corrupted image $\mathbf{g}$ as the input. Then, the model producing the lowest validation error is used for image estimation.

III. Image Estimation Using Piecewise Image Models

Piecewise image models, or PIM's, describe images that obey an image property, such as constancy, linearity, polynomial behavior, or some other more abstract or other specific property on a piecewise basis over the entire image domain. The pieces over which the property holds form a proper partition of the image; each piece is constrained to be of some minimum size (specified by the model degree). The size of a piece may be defined in various ways, such as the minimum dimension along its minor axis. The piecewise model allows for sudden discontinuities in the image property that defines the PIM; there is no explicit discontinuity-detection mechanism, however; the region boundaries naturally evolve as the solution is found.

Two potentially useful piecewise image properties that define PIM's are studied here: *piecewise constancy* (PICO), and *piecewise linearity* (PILI). The associated regression problems defined by (2) are termed *PICO regression* and *PILI regression*, respectively. Both regressions are ill-posed combinatorial problems having nonunique solutions. The corresponding PICO or PILI image estimation problems (6) are easily configured for iterative solution. Naturally, other piecewise models can be defined, such as piecewise quadratic (PIQU) models or higher-order piecewise polynomial models, piecewise exponential (PIEX) models, etc. However, PICO and PILI afford meaningful and simple image descriptions that correspond to commonly encountered natural and synthetic image data, and that adequately demonstrate the framework of

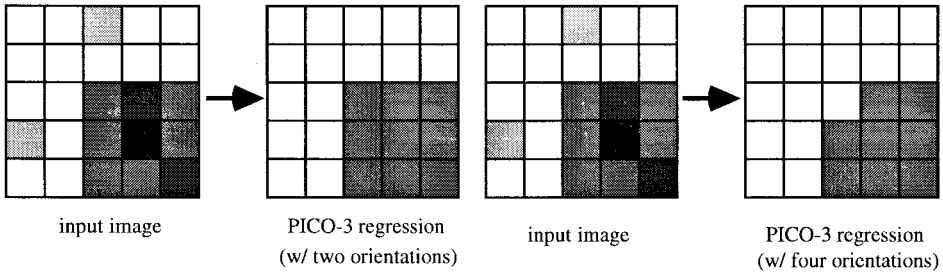


Fig. 1. Illustrative examples of PICO-3 regression using two and four orientations.

this theory. Of course, PICO images define a somewhat more restricted category of imagery; good examples include four-color artwork, printed matter, and binary image data. Another potentially useful application of PICO image estimation is as a preprocessing stage to intensity-based image segmentation. By first forming a PICO image (which defines a coarse segmentation), the segmentation problem is reduced to deciding whether to merge neighboring PICO regions.

The definitions of the PICO and PILI image properties are quite similar, and can be given together as follows.

Definition 1: A one-dimensional (1-D) signal $\mathbf{x}$ is *piecewise constant* (*piecewise linear*) of *degree* m, or PICO-m (PILI-m) if the length of the shortest constant (linear) subsequence in $\mathbf{x}$ is greater than or equal to m.

Thus, each sample is part of a constant (linear) segment of length greater than or equal to m. The lowest degree 1-D PICO (PILI) regression of interest is PICO-2 (PILI-3) , since all signals are PICO-1 (PILI-2).

In defining PILI we make a special dispensation for signals quantized to integer values: the definition is relaxed by allowing each sample to deviate from the nearest real-valued linear trend by no more than unity.

Although PICO and PILI have simple definitions in one dimension, for higher-dimensional signals there is quite a bit of latitude in the definition. The following one supplies an effective piecewise characterization that is also computationally convenient.

Definition 2: A two-dimensional (2-D) image $\mathbf{h}$ is PICO-m (PILI-m) if $\mathbf{h}$ is PICO-m (PILI-m) (in the sense of Definition 1) on every 1-D path along a set of prescribed orientations.

We have experimented with two types of 2-D PICO/PILI definitions: a two-orientation version, and a four-orientation version. The two-orientation PICO (PILI) definition enforces piecewise constancy (linearity) along image columns and rows (linear paths quantized along 90° intervals). The four-orientation definition includes the diagonal orientations (linear paths quantized along 45° intervals).

Four-orientation PICO limits image *streaking*, or highly visible and easy-to-misinterpret constant streaks, similar to those that can occur when a 1-D median filter is applied to an image [6]. Qualitatively, PICO image estimates that utilize the four-orientation constraint exhibit smoother region boundaries, whereas the two-orientation constraint may produce slightly jagged boundaries between the constant regions. There are tradeoffs, of course; imposing PICO along a larger number of orientations creates a more expensive computation of energy in (2) (more paths to check). Also, the four-orientation PICO regression may round corners, as shown in Fig. 1.

In the presence of high-amplitude noise, we have observed that streaking tends to be more severe in image estimates computed under the two-orientation PILI model than under the two-orientation PICO model. In fact, horizontal or vertical streaks can appear along intensity discontinuities.

In the examples, four-orientation PICO and PILI estimates are computed. Although PICO, PILI, and other PIM's share similarities, the assumptions made and the associated implications for implementation differ. These differences will be explored as each model is developed.

A. PICO Image Estimation

If interpreted as an enhancement technique, PICO image estimation successfully accomplishes intraregion smoothing, while preserving important features, especially sharp edges, and removing corruptive noise. As with all PIM's, approximate regressions of different *degrees* are possible, which determines the amount of smoothing.

In (5) and (7), take $PROP = PICO_m$. Then let the set of possible substitutions (of the K possible) for $h(x, y)$ that are members of a piecewise constant vector of length greater than or equal to m in all four orientations be denoted by $\mathbf{pico}_m(x, y)$. Note that only a maximum of eight values must be evaluated to construct $\mathbf{pico}_m(x, y)$, since any piecewise constant solution must be equal to one of the eight neighboring pixels.

Within $\mathbf{pico}_m(x, y)$, the solution with smallest distance to the current value of $h(x, y)$ is assigned to $PICO_m[h(x, y)]$. If the set of local PICO solutions is empty (no local solutions exist), then $PICO_m[h(x, y)]$ is assigned the maximum value $h(x, y) + K - 1$, so the maximum energy penalty is assessed. At each coordinate (x, y), the maximum contribution to (7) is $K - 1$, and the maximum contribution to $E(\mathbf{h})$ is $\lambda(K - 1)$.

There is an interesting relationship between PICO regression or PICO estimation, and a robust class of image-enhancing order statistic filters, known as the weighted majority with minimum range (WMMR) filters [14]. The development of the WMMR filter was motivated by the fact that other impulse-rejecting nonlinear filters, such as the median filter, preserve undesirable monotonic degradation, such as blur, along image edges. The WMMR tends to sharpen such edges by making them more steplike. For a filter window spanning $2m + 1$ samples, the WMMR is implemented by first selecting the $m + 1$ values in the filter window having a minimum range.

The output is computed by a weighted sum of the $m+1$ values. These filters, like the median filter [9], [15], have an interesting root-signal analysis. Indeed, the root signals (signals that remain unchanged by filtering) of a WMMR filter of width $2m+1$ are those signals that are PICO-$(m+1)$. It has also been shown that repeated passes of a WMMR filter eventually produces a PICO root signal. (To achieve a PICO root signal, the WMMR weights must be nonnegative and sum to unity, with unequal first and last weights [14].) We may make the interpretation, then, that the PICO regression directly finds the fixed point of a WMMR filter. This may be stated more strongly: since application of WMMR filters tends to produce PICO results, the *goal* of WMMR filtering may be interpreted as finding a PICO replacement of the input data at the expense of the noise. From this perspective, finding the PICO regression or PICO estimate yields the *best possible* PICO replacement, while the WMMR filter can only deliver a suboptimal one after repeated passes. As will be demonstrated in a numerical comparison later in this section, enhancement results obtained via the WMMR may eliminate important local features that are retained by optimal PICO image estimation.

We note that in [11], a related PICO image estimation procedure was studied. In that work, a *fixed*-size 3×3 neighborhood of every pixel is examined—over each such neighborhood, the target image is assumed constant. Since this is assumed at *every* pixel, this amounts to assuming the image everywhere constant. A penalty is assigned at every pixel by the following strategy: a comparison is made between each pixel in the neighborhood and the pixel under consideration; a penalty of one incurred if unequal, and a penalty of zero if equal. A mean-field annealing algorithm iteratively moderates a tradeoff between minimizing the differences between neighboring pixels and the difference between the original and estimated data. Because of conflict with the data constraint, a PICO image of unknown region scale is obtained. While the PICO constraint developed here leads to a well-defined estimate, the one in [11] is inherently ill defined.

B. PILI Image Estimation

PILI image estimation is also useful for accomplishing intraregion image smoothing without degrading intensity discontinuities. The characteristic set of the associated PILI regression problem is the set of signals that are piecewise linear. Within each image piece, PILI regression allows effective smoothing while retaining intensity trends, which are approximated by linear functions. Thus, the domain of application is broader than afforded by PICO regression/estimation. 1-D PILI regressions were used in [5] to model linear trends in statistical data; piecewise linear topologies for geometric models were explored in [21].

PILI image estimation attempts to enforce linearity on a piecewise basis in a 1-D signal. In 2-D, the PILI vectors effectively form piecewise planar regions. Ideal PILI regressions, when computable (on small-scale problems) retain both step edges and linearly varying ramp edges, while eliminating impulses obtained in a corruptive process. PILI image estimates approximate this behavior, and perhaps, improve upon it. In comparison to PICO regression, PILI estimation yields a more accurate response along slowly varying intensity changes. However, PILI estimates are more difficult to compute and can be less effective than PICO estimation in intense additive noise environments (low SNR, high noise variance), in the sense of image enhancement. The reason for this is that high-amplitude noise processes often continue local groupings of outliers that approximate linear segments; these may be retained or even enhanced by a PILI estimate. However, for lower-intensity noise, the PILI estimates are often very good.

TABLE I
DESCRIPTION OF IMAGERY USED IN PICO AND PILI EXPERIMENTS

Image	Noise Distribution	σ_n	SNR
South Texas	Laplacian	9.5	3.2dB
Mammogram	Uniform	11.1	11.0dB

In (5) and (7), take $PROP = PILI_m$. Denote the set of K possible substitutions for $h(x, y)$ such that a PILI vector of length $\geq m$ is created in all four orientations by $\mathbf{pili}_m(x, y)$. Since the data is discrete, the test for linearity allows for a maximum quantization error of ± 1. The substitution that yields the smallest distance relative to $h(x, y)$ in the set $\mathbf{pili}_m(x, y)$ is assigned to $PILI_m[h(x, y)]$. If $\mathbf{pili}_m(x, y)$ is empty, then $PILI_m[h(x, y)]$ is assigned the maximum value $h(x, y) + K - 1$, yielding the maximum energy.

PILI estimation provides a simple and powerful method for smoothing 1-D signals containing both step edges and ramplike edge transitions. It is also a powerful approach for image enhancement applications, as discrete image data usually contains a proliferation of edge profiles that can be well approximated either by sudden jumps in intensity, or by more gradual linear trends.

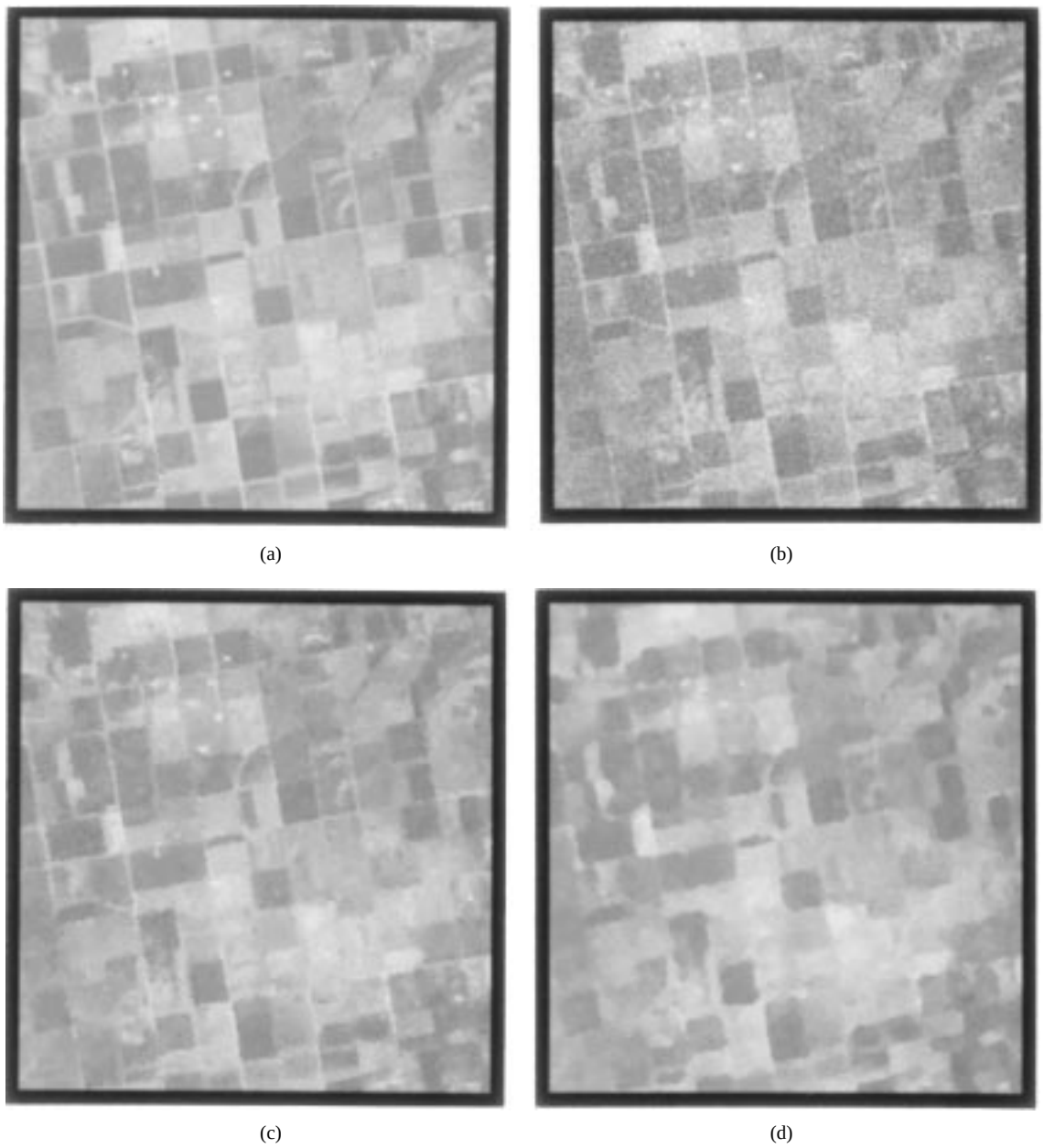
However, 2-D PILI estimation finds a greater degree of computational complexity than might be expected from examination of the 1-D problem. The reason for this is that the strict constraint of piecewise linearity may be difficult to simultaneously satisfy along multiple linear orientations. This leads to poor agreement with the linear model in some locales, which is acceptable, except that some visually misleading local configurations may occur. Conflicts arising between linear paths in the image can result in poor reconstruction of image contours and a failure to eliminate noise. The characteristic properties of simpler models, such as piecewise constancy and (as will be seen) local monotonicity, may be satisfied along several orientations by making single pixel intensity substitutions. By contrast, single pixel changes are often insufficient in satisfying more complicated properties such as with the PILI model.

C. PICO and PILI Image Estimation Examples

In the simulations, we selected images that we deemed to be well approximated by the PIM's, and added noise to them. For these simulations we provide numerical measures of performance expressed in terms of improvement in the error and in the SNR. The SNR of a noisy image is computed via

$$\text{SNR} = 10 \cdot \log_{10}(\sigma_i^2/\sigma_n^2)$$

where σ_i^2 is the variance of the original uncorrupted image and



(a) (b)

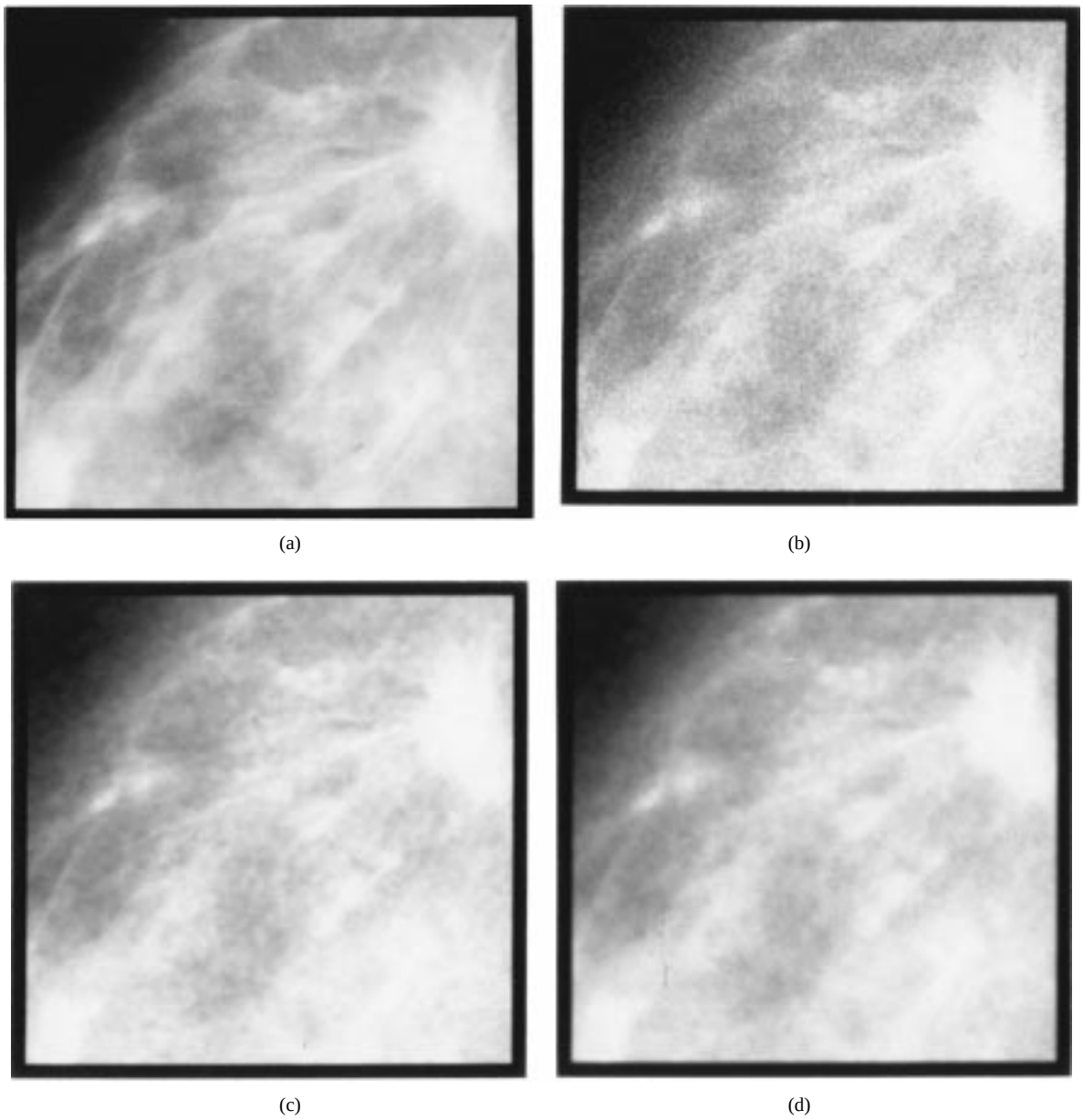
(c) (d)

Fig. 2. PICO image estimation. (a) Original "South Texas" image. (b) Corrupted image. (c) PICO-2 result. (d) WMMR-MED result.

σ_n^2 is the variance of the noise. For each simulation, Table I lists the noise type and noise statistics, and the SNR.

Fig. 2 illustrates PICO image estimation of a 256 $\times$ 256 South Texas SPOT satellite image. Note that the original image Fig. 2(a) is very PICO-like, hence provides an excellent example of the advantage of matching the appropriate image model to the estimation application. In the original SPOT image, several boundaries are ambiguous and noisy outliers are present. The addition of 3.2 dB Laplacian noise creates a nontrivial enhancement problem [Fig. 2(b)]. The PICO image estimate (with l^1-norm data constraint and model constraint) very effectively enacts intraregion smoothing, removing the effects of additive noise while preserving the individual fields, as shown in Fig. 2(c). In terms of region coherence, the PICO-3 image is superior even to the original uncorrupted image, and would be simpler to segment into homogeneous regions. Elimination of noise from cloud cover or from the sensor is an important step in segmenting the agricultural fields shown in the image. Clearly, PICO image estimation can be an effective method for preprocessing noisy images prior to segmentation.

As a comparison, a 5 $\times$ 5 WMMR-MED filter (40 iterations) was also iteratively applied, as shown in Fig. 2(d). This nonlinearly filtered image, while supplying a very PICO-like result, did not retain several of the important features of the image. Using smaller WMMR-MED filters led to severe loss of performance in noise reduction. Although Fig. 2(d) is nearly PICO, several of the South Texas fields are merged together and, in some cases, severely distorted. This blurring effect of the WMMR-MED filter would preclude the possibility of a meaningful segmentation and would also eliminate the possibility of detecting more subtle image regions, such as the roadways separating the fields.



(a) (b) (c) (d)

Fig. 3. PILI image estimation. (a) Mammogram image. (b) Corrupted image. (c) PILI-4 result. (d) Length-5 Δ-OS result.

Fig. 3 is an example of PILI image estimation. In this case, a 256 $\times$ 256 digital x-ray mammogram, Fig. 3(a), is processed. The image was selected since it is composed largely of fairly smooth regions with few abrupt transitions. A uniform-noise corrupted image is shown in Fig. 3(b). The resulting PILI image estimate, using the l^∞-norm in the data constraint, [Fig. 3(c)] is nicely smoothed, but also retains the important features of the original image. Here, nonlinear image estimation with respect to the PILI-4 characteristic set produced an improvement in the SNR of 8.5 dB. If PICO image estimation were employed instead, it is likely that misleading false contours would have developed in the solution image, thus distorting possible interpretation of the parenchymal tissues revealed in the mammogram.

As a filter comparison to the PILI image estimation method, we applied a specific *order statistic* (OS) filter to the noisy mammogram [7]. Within a finite window, the filter algebraically orders the intensities within the window, then linearly weights them using a piecewise linear (triangular) weighting to compute the output. Thus, the filter, called the Δ-*OS filter* (triangle OS filter) was selected, since it is near optimal for heavy-tailed noise in minimum variance sense [7]; it is highly robust, and it supplies a linear weighting to naturally ordered samples near intensity transitions. This makes it a fair comparison for a piecewise linear fit. It was implemented by applying a 1-D Δ-OS filter along both the rows and columns of

TABLE II
PICO AND PILI IMAGE ESTIMATION AND FILTERING RESULTS

Image	Method	MSE	ΔSNR
South Texas	PICO-2 Regression	22.6	+5.75dB
South Texas	WMMR-MED (40x)	31.4	+4.32dB
Mammogram	PILI-4 Regression	16.3	+8.50dB
Mammogram	Δ-OS Filter	43.0	+4.29dB

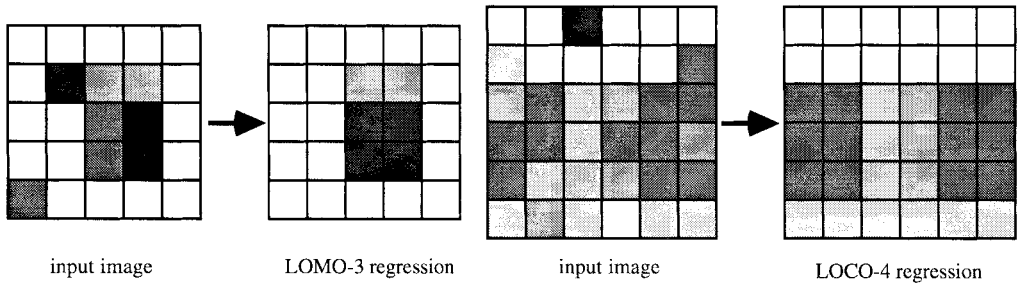


Fig. 4. Illustrative examples of LOMO-3 and LOCO-4 regressions in two dimensions.

the image (a common strategy that agrees with the row/column definition of PILI utilized here). A window span of five was used in the Δ-OS filter example shown in Fig. 3(d)—the smallest possible Δ-OS filter, since the length-3 Δ-OS filter is equivalent to the length 3 median filter. Larger window sizes and multiple iterations resulted in inferior (blurred) results. Although the result in Fig. 3(d) is fairly smooth, several important features (including the possible tumor!) have been eliminated. By comparison, the PILI-4 estimate exhibits superior feature preservation while still effectively smoothing the noise.

Table II gives numerical results for each nonlinear estimation method, showing the MSE with respect to the original uncorrupted image and the improvement in SNR from the corrupted image, which is given by

$$\Delta\text{SNR} = 10 \log_{10} ||\mathbf{i} - \mathbf{g}||^2 / ||\hat{\mathbf{g}} - \mathbf{g}||^2.$$

It can be seen that in each case, the mean square error (MSE) was substantially smaller using the nonlinear estimator. The improvement in SNR was also better, sometimes dramatically so.

IV. Image Estimation Using Local Image Models

The second class of image models studied, *local image models* (LIM's), describe images that obey an image property, such as monotonicity, convexity/concavity, or other specific property over *every* image region of specified size and geometry. Because the characteristic properties are required to hold everywhere, LIM's require more flexible image properties than do PIM's; for example, the only images that are locally constant everywhere are globally constant; the only images that are locally linear everywhere are also globally linear. Thus constancy and linearity are image properties that do not lead to interesting LIM's. By contrast, piecewise monotonic regressions/estimates and piecewise concave/convex regressions/estimates lead to viable models.

Since they are required to hold everywhere, the characteristic properties of LIM's must have the ability to capture a broad range of image structures. Two such characteristic properties are studied here: *local monotonicity* (LOMO), which defines images that are monotonic on every local region of specific geometry, and *local convexity/concavity* (LOCO), which defines images that are convex or concave on every local region of specific geometry. We refer to the associated regression problems (2) as *LOMO regression* and *LOCO regression*, and the estimation problem (6) by *LOMO image estimation* and *LOCO image estimation.*

Again, the size and shape of the local geometry over which the characteristic property is constrained to hold is an important specification, and is part of the definition of a LIM. The definitions of LOMO and LOCO signals in both 1-D and 2-D are again quite similar, and given together, as follows.

Definition 3: A 1-D signal $\mathbf{x}$ is LOMO-m (LOCO-m) if every subsequence of $\mathbf{x}$ of length $\leq m$ is monotonic (is either convex or concave).

Note that since every 1-D signal is LOMO-2 (LOCO-3), LOMO-3 (LOCO-4) is the smallest property degree of interest.

Definition 4: A 2-D image $\mathbf{h}$ is LOMO-m (LOCO-m) if $\mathbf{h}$ is monotonic (is either convex or concave) on every 1-D path of length $\leq m$ along a set of prescribed orientations.

Both two- and four-orientation LOMO and LOCO versions have been tested; the differences in solution quality between two-orientation and four-orientation implementations was found to generally be quite small; indeed, image streaking appears not to be a problem with LIM's, at least those tested thus far. Therefore, the less expensive two-orientation version was used exclusively in the LOMO and LOCO examples (see Fig. 4).

A. LOMO Image Estimation

The smoothing properties of locally monotonic (LOMO) regression have previously been studied in some depth for 1-D signals in [18], [19]. Local monotonicity is well suited for describing images, since the model embodies image structures that include step edges, ramp edges, and all types of monotonic edge profiles. The LOMO model also captures smoothness. Thus, LOMO image estimates tend to have well-preserved edges and effectively smoothed noise.

Now take $PROP = LOMO_m$ in (5) and (7). The set of possible substitutions (of the K possible) for $h(x, y)$ such that $h(x, y)$ is a member of a LOMO segment of length $\geq m$ along each prescribed orientations is denoted $\mathbf{lomo}_m(x, y)$. Within this set, the solution having the smallest distance to the current value of $h(x, y)$ is $LOMO_m[h(x, y)]$. If $\mathbf{lomo}_m(x, y) = \phi$, then set $LOMO_m[h(x, y)] = h(x, y) + K - 1$.

Just as PICO regression and PICO image estimation are related to the WMMR nonlinear enhancement filter (through sharing of fixed points), the techniques of LOMO regression and LOMO image estimation are related to the *median filter*. Indeed, it was research into the interesting properties of the median filter that first led to the introduction of the concept

of locally monotonic regression [18]. Just as the PICO signals are the fixed points of the WMMR filters, LOMO signals are the fixed points of median filters (with a well-established 1-D fixed point theory [9], [15]). Similar arguments may be made in favor of LOMO regression and LOMO image estimation as were made for PICO-based methods. Since repeated filtering with a median filter leads inevitably to a LOMO signal, then median filtering may be seen as a method for inducing "LOMOness" on a signal. LOMO regression/estimation is also such a technique; however, with a more directed goal of finding a *best* LOMO signal. As might be expected, there are similarities between median filtering results and LOMO estimation results, as will be seen in the simulation.

B. LOCO Image Estimation

LOCO regression for 1-D signals was first studied in [19]. The idea behind locally convex/concave (LOCO) image estimation is that a signal can be smoothed by limiting the rate of change in monotonicity within every signal region. This is a very novel measure of signal smoothness, and certainly, LOCO regression/estimation is somewhat specialized. For example, LOCO regression does not adequately preserve step edges. Also, the LOCO model constraint is not particularly effective at eliminating large noise impulses; undesirable LOCO oscillations may be created on the image surfaces. However, for images that contain smoothly changing edge structures, or LOCO oscillatory patterns, the approach can be very effective.

This time, take $PROP = LOCO_m$ in (5) and (7). Let $\mathbf{loco}_m(x, y)$ be the set of possible solutions for $h(x, y)$ that are members of locally convex/concave segments of length $\geq m$ along both the vertical and horizontal orientations. The member of $\mathbf{loco}_m(x, y)$ having the smallest distance to the current value of $h(x, y)$ is the value of $LOCO_m[h(x, y)]$. If $\mathbf{loco}_m(x, y)$ is empty, then $LOCO_m[h(x, y)]$ is assigned the maximum value $h(x, y) + K - 1$.

Like piecewise linearity, the constraint for local convexity/concavity is expensive to compute, since several nontrivial solutions to the LOCO constraint may exist at each pixel location. However, unlike PILI regression/estimation, satisfying the LOCO property locally in two directions is not difficult when using single pixel changes at each iteration of an optimization routine.

As a method of image enhancement, LOCO image estimation has not been previously applied to real-world image data. Since the LOCO model does not preserve step edges, the domain of application is somewhat limited, and certainly would preclude images of most man-made, indoor scenes. Nevertheless, LOCO image estimation may used efficaciously in specific image applications, as well as in extended domains such as smoothing of nonabrupt audio signals immersed in noise, or for enhancing other inherently bandlimited (lowpass) signals.

C. LOMO and LOCO Image Estimation Examples

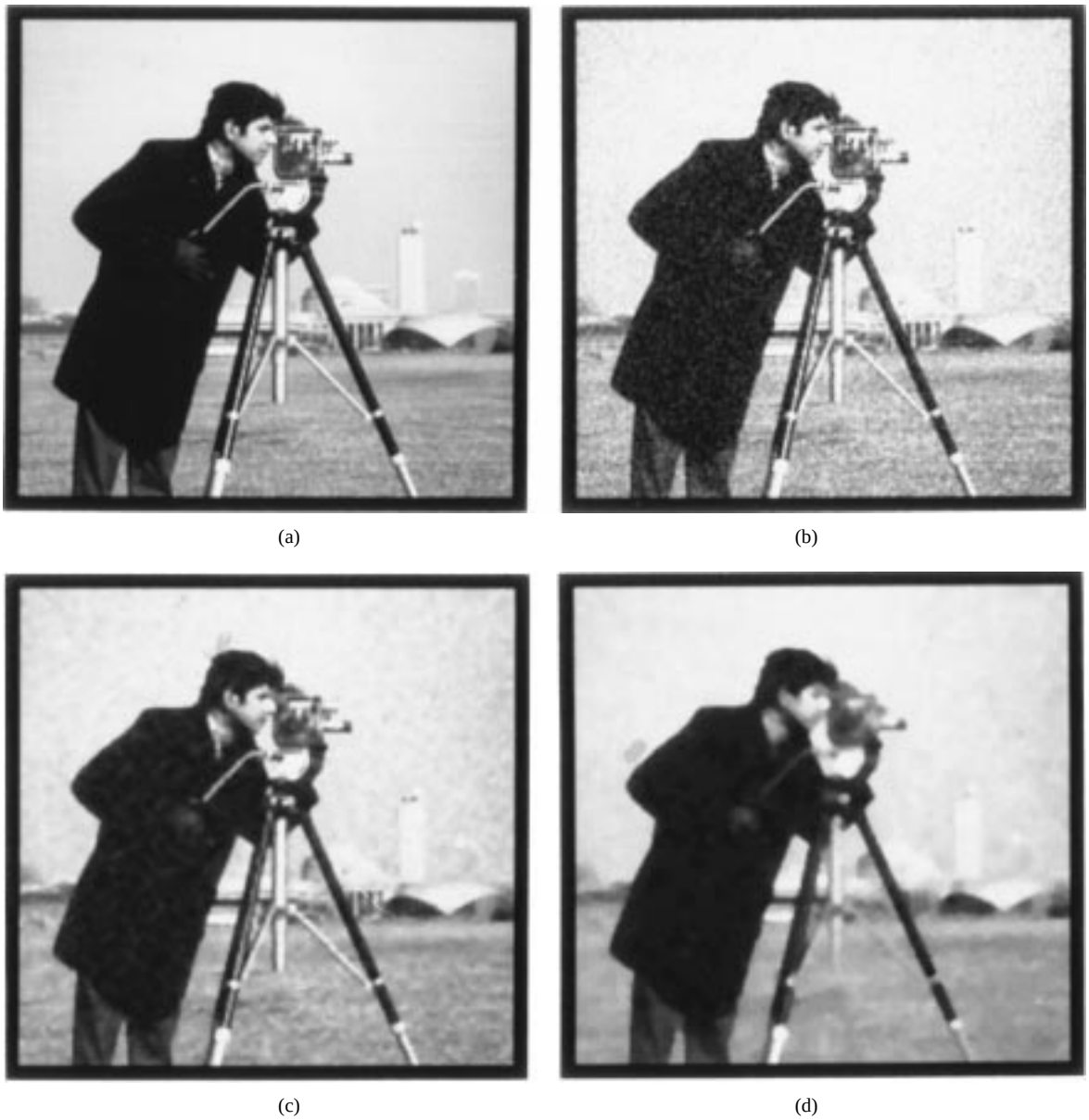
In each simulation we attempt to utilize, for each estimation method, an input image that is effectively modeled by the appropriate LIM. Table III lists the relevant input image statistics.

TABLE III
DESCRIPTION OF IMAGERY USED IN LOMO AND LOCO EXPERIMENTS

Image	Noise Distribution	σ_n	SNR
Cameraman	Gaussian	19.0	10.0dB
Tree	Laplacian	28.1	1.7dB

Next, Fig. 5 depicts filtering of the cameraman image [Fig. 5(a)], containing a mixture of detailed and smooth regions. This image was selected since the LOMO model is intended to be quite generic. A Gaussian-noise corrupted version of this image was created, as shown in Fig. 5(b); hence, the data constraint was defined using the l^2-norm. Fig. 5(c) shows the result of LOMO-3 image estimation. The flexibility of the LOMO model is evident—through the simultaneous smoothing of large-scale regions such as the background, and the retention of the finely detailed features such as the cameraman's facial features. Notice the smooth contours and the natural ramplike edges, such as the shading on the tripod. Several small but physically meaningful regions in the image, such as the eyes and the individual camera components, are retained in the LOMO image estimate. By comparison, the rootlike signal generated by successive application (40 iterations) of a 3×3 square window median filter [Fig. 5(d)] is quite smooth in the global sense, but at the loss of detail, and the creation of several unattractive blotchy patches [6]. Note also the blurring of facial features, the camera, and the buildings in the background.

Finally, Fig. 6 depicts LOCO estimation of a severely corrupted (1.7 dB) image of a tree's cross section. The application of the LOCO image model is appropriate, because the "tree" image [Fig. 6(a)] exhibits an approximately sinusoidally varying intensity pattern—and few steplike edges. The noisy image, which was corrupted with Laplacian-distributed noise, is severely degraded [Fig. 6(b)]. However, the LOCO image estimate (defined using the l^1-norm for the data constraint) shown in Fig. 6(c) is a very smooth result that corresponds very well with the intensity profile of the original image in Fig. 6(a). As a method of comparison with an appropriate nonlinear filter, a *moving LOCO* filter was applied to the image. The moving LOCO filter, defined here for the first time, forces the digital signal to be locally convex/concave, in the 1-D sense, along the row and columns of the image. Note that a sampled locally convex/concave signal has a difference signal $(x_i - x_{i-1}, x_{i+1} - x_i, x_{i+2} - x_{i+1}, \cdots)$ that is locally monotonic. Specifically, a 1-D LOCO-$m + 1$ signal has an associated difference signal that is LOMO-m. Therefore, a 1-D LOCO signal can be computed by forcing the difference signal to be LOMO. This is accomplished by first computing the difference signal along an image row or column (discrete differentiation), using the moving LOMO filter defined in [19] to create a LOMO difference signal, then summing the new differences (discrete integration) to compute the LOCO signal. This operation is applied to each image row and column. Note that this operation does not guarantee that the result will be LOCO in the 2-D sense. However, it has the advantage of



(a) (b) (c) (d)

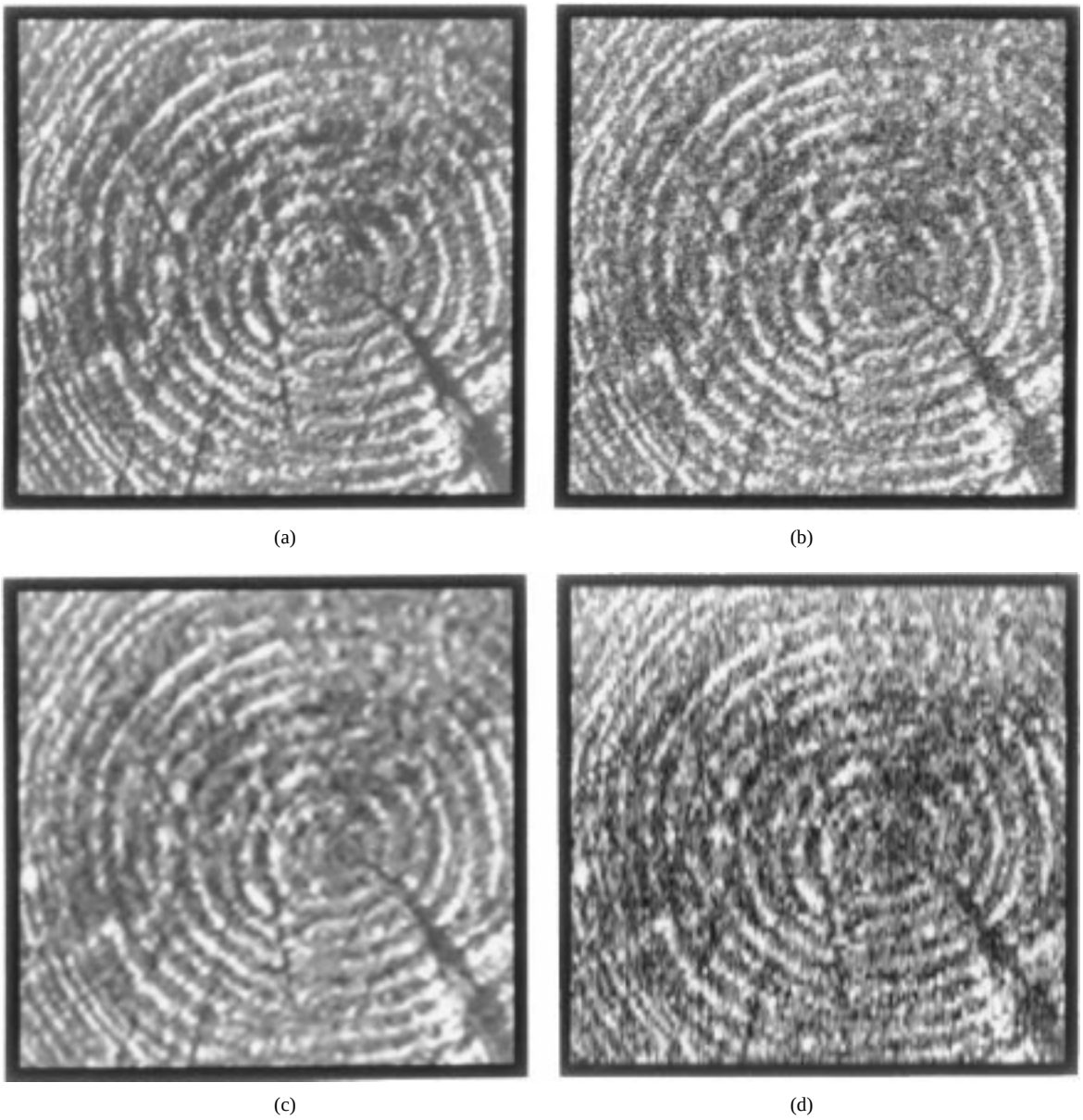
Fig. 5. LOMO image estimation: (a) Cameraman image. (b) Corrupted image. (c) LOMO-3 result. (d) Iterated median filter result.

speed. The moving LOCO filter result is shown in Fig. 6(d). Although the result is reasonable, this filter fails to match the smooth, high-quality result of the LOCO-4 image estimate.

Table IV lists the errors incurred by both LIM-based nonlinear estimation and by the comparative nonlinear filters used. In each case, the MSE was again substantially smaller using the nonlinear estimator. Accordingly, the improvement in SNR was also superior.

V. Iterative Solution via GDA

The nonlinear image estimation problems studied here are all *combinatorial*, *multistate* (full gray level), and *nonconvex*. Combinatorial optimization problems have discrete, finite solution spaces that increase exponentially (equivalently, as $N!$) as the problem size increases [16]. Clearly, the image estimation problem is combinatorial as the number of possible solutions increases as K^{N^2}, where K is the number of possible pixel intensities and N^2 is the number of pixels in the image. The estimation problem is, of course, inherently multistate (as opposed to binary). In the examples presented here, 8-b data is used so that each optimization variable has 256 discrete states. The energy functions defined for the PICO, PILI, LOMO, and LOCO models are nonconvex; hence, globally optimal solutions cannot be found using steepest descent (local search). Suboptimal local minima can be avoided through the statistical hill climbing of stochastic simulated annealing (SA). However, even "practical" implementations of SA have an unrealistic computational expense for gray-level image estimation applications. As an effective alternative, we formulate solutions to nonlinear estimation problems using *generalized deterministic annealing* (GDA), a very recent optimization technique that provides high-quality solutions for time-critical applications [2]. Unlike previous optimization methods used in image processing applications, GDA is a general-purpose tool



(a) (b) (c) (d)

Fig. 6. LOCO image estimation: (a) Tree image. (b) Corrupted image. (c) LOCO-4 result. (d) Moving LOCO filter result.

TABLE IV
LOMO/LOCO IMAGE ESTIMATION AND FILTERING RESULTS

Image	Method	MSE	ΔSNR
Cameraman	LOMO-3 Regression	137.2	+3.62dB
Cameraman	Median Filter (40x)	251.9	+0.98dB
Tree	LOCO-4 Regression	165.4	+5.76dB
Tree	Moving LOCO Filter	1048.3	-2.26dB

for multistate problems, is characterized by rapid, guaranteed convergence and by the ability to escape undesirable local solutions. In contrast to SA, GDA can easily be implemented in a true parallel fashion on a single instruction multiple data (SIMD) architecture, without the need for "divide and conquer" schemes.

GDA directly estimates the limiting solution of the SA algorithm. The iterative, stochastic solution of SA may be modeled as a Markov process [1]. Each state in the Markov chain represents a specific, unique solution. For an optimization problem with N^2 variables with K possible states, the SA Markov chain has K^{N^2} possible states. Solution changes occur according to the SA transition probabilities. At each temperature in the annealing process, the chain converges to an equilibrium state (stationary distribution) after many transitions. At high temperatures, the stationary distribution is uniform, where all solutions in the chain have equal probability. As the temperature is slowly reduced in the annealing process, the chain *freezes* into a globally optimal solution. To directly estimate the limiting solution of the SA algorithm, GDA utilizes N^2 separate *local Markov chains* of length K. Each local Markov chain represents the state of an optimization variable (e.g., pixel intensity). Using the SA transition probabilities, GDA iteratively computes the distribution (not the state) of each local Markov chain at a given temperature. Due to the shorter K-length of the

GDA local Markov chains, the equilibrium state at each temperature is achieved quickly after only a few iterations. As the temperature is lowered in the annealing process, a single solution emerges for each optimization variable (the distribution becomes singular at the final state for the optimization variable). When all of the N local chains have become frozen at a final state, the estimate of the optimal solution is completed. The approximate solution corresponds to a local minimum in the energy function [2].

Denote the distribution for the local Markov chain for the pixel intensity $h(x, y)$ at iteration t by $\pi^t(x, y) = [\pi^t(x, y; 0)\pi^t(x, y; 1)\cdots\pi^t(x, y; K-1)]$. The ith component $\pi^t(x, y; \bullet)$ is the probability mass function of $h(x, y)$ at iteration t. The K components of $\pi^t(x, y)$ are the K states of the local Markov chain that correspond to the K possible intensities for a given pixel. At each iteration, a new density is computed for each pixel intensity based on the previous distribution. An update for the ith component at (x, y) is accomplished by

$$\pi^{t+1}(x, y; j) = (1/K) \sum_{i=0}^{K-1} A_T(x, y; i, j) \cdot [\pi^t(x, y; i) + \pi^t(x, y; j)] \tag{8}$$

where

$$A_T(x, y; i, j) = 1/[1 + \exp((1/T)\{E_{loc(x,y)}[h(x, y) = j] - E_{loc(x,y)}[h(x, y) = i]\})] \tag{9}$$

and $E_{loc(x,y)}[h(x, y) = i]$ is the *local energy* at (x, y) when pixel $h(x, y)$ is assigned a value of i. The local energy is computed using the mean field estimates of neighboring variables. The mean field estimate of the pixel value $h(x, y)$ at time t is

$$h^*(x, y) = \text{int}\left[\sum_{i=0}^{K-1} i\pi^t(x, y; i)\right] \tag{10}$$

where int$(\bullet)$ is the nearest integer function. Uniform convergence for the estimate may be described as $\nabla h^*(x, y) < \delta \forall x, y : 0 \le x \le N-1, 0 \le y \le N-1$. $\nabla h^*(x, y)$ is the change in $h^*(x, y)$ between successive iterations. For the probability densities, uniform convergence may be stated as

$$|\pi^{t+1}(x, y, i) - \pi^t(x, y, i)| < \varepsilon \qquad \forall x, y, i: 0 \le x \le N-1, 0 \le y \le N-1, 0 \le i \le K-1. \tag{11}$$

For image estimation, $\varepsilon = 1/K^2$ guarantees that the changes in pixel intensity (11) are less than unity $[\nabla h^*(x, y) < 1]$. The number of iterations needed to obtain this measure of uniform convergence at a temperature T is given by [2]

$$\eta_c(T) = \frac{\ln\left(\frac{1}{K^2}\right)}{\ln\left[\left(\frac{K-2}{2K}\right)e^{-T/\Delta E_{\max}} + \frac{1}{2}\right]} \tag{12}$$

where T_0 is the initial annealing temperature T_f is the final temperature. Using the guidelines in [2] for the image estimation problem

$$T_f = \Delta E_{\min}/\ln[K^2 - K - 1] \tag{13}$$

and

$$T_0 = \Delta E_{\max}/\ln\left[\frac{K}{K-2}\right] \tag{14}$$

where $\Delta E_{\max}$ is the maximum energy change, and $\Delta E_{\min}$ is the minimum (nonzero) energy change possible with one variable change. Because the minimum and maximum energy changes depend on the realization of $\mathbf{PROP}(\mathbf{h})$, $\Delta E_{\max}$, and $\Delta E_{\min}$ must be computed for each PIM and LIM. For the PICO, PILI, LOMO, and LOCO models, assuming integer-valued pixel intensities, define the minimum p-semimetric value in (4) as γ where

$$\gamma \equiv \min \|\mathbf{g} - \mathbf{h}\|_D = \min \|\mathbf{h} - \mathbf{PROP}(\mathbf{h})\|_M = (K-1)[N^{2/p} - (N^2-1)^{1/p}] \tag{15}$$

for $p \ge 1$. Therefore

$$\Delta E_{\min} = \min\{\gamma, \lambda\gamma\}. \tag{16}$$

Since each of the four PIM's and LIM's have the same maximum contribution to the energy functional

$$\Delta E_{\max} = \max\{K-1, \lambda(K-1)\}. \tag{17}$$

An effective implementation of GDA for the nonlinear image estimation problem follows.

Step 1. *Initialization:* Set $T = T_0$ and set $\pi^0(x, y; i) = 1/K \forall x, y, i : 0 \le x \le N-1, 0 \le y \le N-1, 0 \le i \le K-1$.

Step 2. *Iteration:* Use (8) to update $\pi^t(x, y; i) \forall x, y, i : 0 \le x \le N-1, 0 \le y \le N-1, 0 \le i \le K-1$.

Step 3. *Equilibrium:* If the number of iterations at the current temperature, $n > n_c(T)$, then set $T = \tau T$ (where $0.9 \le \tau \le 0.95$).

Step 4. *Saturation:* If $T < T_f$, stop. Else, return to Step 2.

Additional speedup may be obtained using *windowed* GDA (WGDA) [2], where only a small window of states of length K_w in the K-length local chains are active at any time. The window is centered at the mean field estimate (10) for each pixel; window shifts are limited to one state/iteration to prevent oscillations. In all the examples presented here, a WGDA implementation with $K_w = 9$ was utilized. The WGDA affords over two orders of magnitude of improvement in speed over a practical SA algorithm, for comparable solution quality.

VI. Concluding Remarks

The characteristic set—e.g., PICO, PILI, LOMO, LOCO—defines the image model used. Naturally, the model used must be appropriate. An image that was originally or nearly LOMO is an ideal candidate for LOMO estimation. However, generalizations can be made. For images of man-made environments, PICO, PILI, and LOMO are quite tenable models, since they all effectively preserve steplike edges that are usually numerous in man-made scenes. For images of synthetic environments containing surfaces having uniform reflectance profiles (e.g., a robotics application),

PICO estimation is quite powerful. Natural scenes contain a combination of sharp steplike edges and gradually changing ramplike edges, so LOMO estimation is an excellent choice. PILI estimation displays superior performance on many images with smooth intensity profiles, but at greater expense. The applications for LOCO image estimation are much more restricted. One application might be estimating 2-D sinusoidal gratings. The PIM's and LIM's considered here do provide a diversity of image models for image estimation tasks, although, no doubt, many others can be defined.

We are currently studying application of PIM's and LIM's as set-theoretic constraints on the restoration of images that have been both blurred and corrupted with noise. The extension of the PIM's and LIM's to color and multispectral imagery is still open. Currently, the image estimation process using the piecewise and local models could be applied to each spectral band independently. The development of PIM's and LIM's that incorporate information from several spectral bands simultaneously could be useful to the color imaging and to the remote sensing community.

REFERENCES

[1] E. H. L. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing.* New York: Wiley, 1987.

[2] S. T. Acton and A. C. Bovik, "Generalized deterministic annealing," *IEEE Trans. Neural Networks,* vol. 7, pp. 686–699, 1996.

[3] H. C. Andrews and B. R. Hunt, *Digital Image Restoration.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

[4] A. Blake and A. Zisserman, *Visual Reconstruction.* Cambridge, MA: MIT Press, 1987.

[5] F. L. Bookstein, "On a form of piecewise linear regression," *Amer. Stat.,* vol. 29, pp. 116–117, 1975.

[6] A. C. Bovik, "Streaking in median filtered images," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-35, pp. 493–503, 1987.

[7] A. C. Bovik, T. S. Huang, and D. C. Munson, "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-31, pp. 1342–1350, 1983.

[8] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Processing,* vol. 1, pp. 322–336, 1992.

[9] N. C. Gallagher and G. L. Wise, "A theoretical analysis of the properties of median filters," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-29, pp. 1136–1141, 1981.

[10] D. Geman and S. Geman, "Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-6, pp. 721–741, 1984.

[11] H. Hiriyannaiah, G. Bilbro, W. Snyder, and R. C. Mann, "Restoration of piecewise-constant images by mean-field annealing," *J. Opt. Soc. Amer.,* vol. 6, 1989.

[12] F. C. Jeng and J. W. Woods, "Image estimation by stochastic relaxation in the compound Gaussian case," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing,* New York, NY, Apr. 1998, pp. 1016–1019.

[13] A. K. Katsaggelos, Ed., *Digital Image Restoration.* Berlin, Germany: Springer-Verlag, 1991.

[14] H. G. Longbotham and D. Eberly, "The WMMR filters: A class of robust edge enhancers," *IEEE Trans. Signal Processing,* vol. 41, pp. 1680–1684, 1993.

[15] H. G. Longbotham and A. C. Bovik, "Theory of order statistic filters and their relationship to linear FIR filters," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 37, pp. 275–287, 1989.

[16] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity.* Englewood Cliffs, NJ: Prentice-Hall, 1982.

[17] S. J. Reeves and R. M. Mersereau, "Automatic assessment of constraint sets in image restoration," *IEEE Trans. Image Processing,* vol. 1, pp. 119–122, 1992.

[18] A. Restrepo (Palacios) and A. C. Bovik, "Locally monotonic regression," *IEEE Trans. Signal Processing,* vol. 41, pp. 2796–2810, 1993.

[19] A. Restrepo (Palacios), "Locally monotonic regression and related techniques for signal smoothing and shaping," Ph.D. dissertation, Univ. Texas, Austin, May 1990.

[20] A. Restrepo (Palacios) and A. C. Bovik, "On the statistical optimality of locally monotonic regression," *IEEE Trans. Signal Processing,* vol. 42, pp. 1548–1550, 1994.

[21] C. P. Rourke and B. J. Sanderson, *Introduction to Piecewise Linear Topology.* Berlin, Germany: Springer-Verlag, 1981.

[22] N. D. Sidiropoulos, "Fast locally monotonic regression," *IEEE Trans. Signal Processing,* vol. 45, pp. 389–395, 1997.

[23] T. Simchony, R. Chellapa, and Z. Lichtenstein, "Graduated nonconvexity algorithm for image estimation using compound Gauss Markov field models," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing,* Glasgow, U.K., 1989, pp. 1417–1420.

Scott T. Acton (S'89–M'93) received the B.S. degree in electrical engineering from Virginia Polytechnic Institute and State University, Blacksburg, in 1988, and the M.S. and Ph.D. degrees in electrical engineering from the University of Texas, Austin, in 1990 and 1993, respectively.

He has worked in industry for AT&T, the MITRE Corporation, and Motorola, Inc. Currently, he is an Associate Professor in the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, where he directs the Oklahoma Imaging Laboratory. The laboratory is sponsored by several organizations, including the Army Research Office, NASA, and Lucent Technologies. His research interests include multiscale image representations, diffusion algorithms, image morphology, and image restoration.

Dr. Acton is an active participant in the IEEE, ASEE, SPIE, and Eta Kappa Nu. He is the winner of the 1996 Eta Kappa Nu Outstanding Young Electrical Engineer Award, a national award that has been given annually since 1936. Locally, he has been selected as the 1997 Halliburton Outstanding Young Faculty Member.

Alan C. Bovik (S'80–M'81–SM'89–F'96) was born in Kirkwood, MO, on June 25, 1958. He received the B.S. degree in computer engineering in 1980, and the M.S. and Ph.D. degrees in electrical and computer engineering in 1982 and 1984, respectively, all from the University of Illinois, Urbana-Champaign.

He is currently the General Dynamics Endowed Fellow and Professor in the Department of Electrical and Computer Engineering, the Department of Computer Sciences, and the Biomedical Engineering Program, University of Texas, Austin, where he is also the Associate Director of the Center for Vision and Image Sciences. During the Spring of 1992, he held a visiting position in the Division of Applied Sciences, Harvard University, Cambridge, MA. His current research interests include digital video, image processing, computer vision, wavelets, 3-D microscopy, and computational aspects of biological visual perception. He has published more than 250 technical articles in these areas and holds U.S. patents for the image and video compression algorithms VPIC and VPISC. He is a Registered Professional Engineer in the State of Texas and is a frequent consultant to industry and academic institutions.

Dr. Bovik participates in a wide range of professional activities. He is on the Board of Governors, IEEE Signal Processing Society, since 1996. He is Editor-in-Chief of the IEEE TRANSACTIONS ON IMAGE PROCESSING, also since 1996. He has been an Associate Editor of IEEE SIGNAL PROCESSING LETTERS (1993–1995) and an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING (1989–1993). He has been on the Editorial Board of the *Journal of Visual Communication and Image Representation* (1992–1995). He currently serves on the Editorial Board of *Pattern Recognition* (since 1988), *Pattern Analysis and Applications* (since 1997), and is Area Editor of *Graphical Models and Image Processing* (since 1995). He was on the Steering Committee of IEEE TRANSACTIONS ON IMAGE PROCESSING (1991–1995); was the Founding General Chairman of the First IEEE International Conference on Image Processing, Austin, TX, 1994; and served as Program Chairman of the SPIE/SPSE Symposium on Electronic Imaging 1990. He is a winner of the University of Texas Engineering Foundation Halliburton Award and a two-time Honorable Mention winner of the International Pattern Recognition Society Award.