

An object detection strategy for uncooled infrared imagery

H. A. SCHMITT*, J. G. RIDDLE and T. M. BRUCKS

Raytheon Company, Missile Systems, 1151 E. Herman Road, Building M09, M/S 3, Tucson, AZ 85706, USA; e-mail: haschmitt@west.raytheon.com

R. R. COIFMAN and I. COHEN

Fast Mathematical Algorithms & Hardware, 1020 Sherman Avenue, Hamden, CT 06514, USA

(Received 15 February 2002; revision received 20 April 2002)

Abstract. An overall strategy for the very difficult problem of object detection using uncooled infrared (UCIR) sensors is discussed. The UCIR sensors are based on micro-bolometer technology and thus differ significantly from cooled infrared sensors that employ photon-counting detectors. As such, UCIR imagery tends to be very low contrast, since the sensor operates over a broad spectral band; and blurry, because of the long integration times. Ideally, the UCIR imagery would be preprocessed using an appropriate image reconstruction/restoration algorithm. If the sources of image degradation are understood and lend themselves to accurate modelling, the image reconstruction can be solved as an inverse problem. Most often this is not the case and the problem is solved using minimization approaches, such as blind deconvolution. Because image reconstruction/restoration approaches tend to be very throughput intensive, they are rarely performed in a tactical environment. More typically, a detection algorithm is applied directly to the UCIR imagery. In this paper, Local Singular Value Decomposition (LSVD) is evaluated for anomaly detection. LSVD uses local statistics to identify anomalous regions and is very good at identifying local texture differences; it appears to work quite well on UCIR imagery. Target detection results are presented for a simulated data set.

1. Introduction

Infrared imaging sensors that operate without cryogenic cooling have the potential to provide the military user with exceptional night vision capabilities packaged in a device of extremely small size, weight, and power. This would significantly reduce the cost and accelerate the implementation of sensors for applications such as targeting, surveillance, and threat warning. However, the performance of uncooled infrared sensors is still inferior to that of cooled sensors. This performance gap limits the number of applications and precludes the widespread use of uncooled infrared sensors in military missions. While uncooled sensors are being considered as replacements for cooled sensors in some applications, perhaps more importantly, the unique characteristics of the uncooled

* Author for correspondence.

sensors are spawning novel uses of the technology. Very small, low-power sensors with moderate levels of performance are possible with the uncooled infrared technology. In addition, applications such as micro-vehicles and robotics demand an extremely lightweight imaging sensor.

There is currently a great deal of interest in uncooled IR sensors within the Department of Defense community for Automatic Target Acquisition (ATA) for smart munitions. The obvious attraction of uncooled IR sensors over the more traditional cooled sensors is their low cost. This cost advantage should become even more pronounced with the economies-of-scale expected from commercial applications, particularly the automotive industry. The trade-off for achieving this greatly reduced cost is degradation in image quality that places a significantly greater burden on the ATA algorithms. There are a number of very exciting new approaches, such as multiplexed imaging [1–3], currently being investigated to improve the performance of uncooled sensors; however, they are not mature enough to be implemented in current generations of munitions.

The remainder of this manuscript is structured as follows. Section 2 briefly describes Raytheon's UCIR sensor. We discuss in section 3 the UCIR imagery that was available for algorithm evaluation. This imagery consists of both high fidelity simulation data as well as data collected during a recent captive flight test (CFT). Image reconstruction is introduced in section 4 and a specific example is provided. An overview of the Local Singular Value Decomposition (LSVD) anomaly detection algorithm is given in section 5. The results from processing a set of simulated UCIR imagery is presented in section 6. Finally, we summarize our results and conclusions in section 7, which also contains a discussion of further research.

2. Description of UCIR sensor

Raytheon's uncooled IR camera is based on the SB-246 detector array developed by Santa Barbara Research Centre. The SB-246 is a 320×240 room-temperature focal plane array (FPA). This FPA is based upon silicon microbolometer technology. The bolometer material is vanadium oxide (VOx). A picture of the FPA is shown in figure 1. The FPA specifications are detailed in table 1.

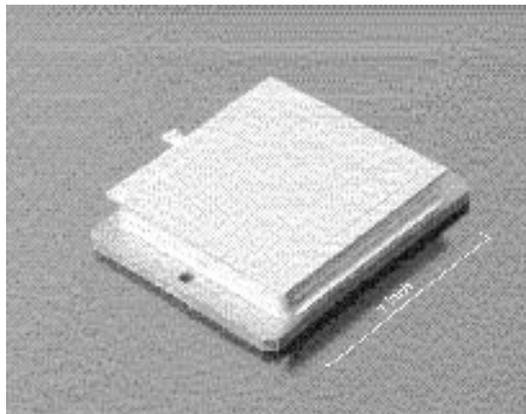


Figure 1. SB-246 focal plane array.

Table 1. UCIR sensor specifications.

Spectral band	8–12 μm	Array size	320 \times 240
IFOV	0.22 mrad	Frame rate	\leq 60 Hz
Pitch	25 μm	NETD	80 mK (f/1 aperture)
Non-uniformities	\leq 100 mV (peak-to-peak)		

3. Description of test imagery

For the processing described in the current paper, we use Raytheon’s modelling environment—simulation technology image generation (STIG)—that was specifically developed for uncooled IR scene generation. STIG is a high fidelity modelling tool that uses a collection of measured backgrounds, a series of turntable target data, a detailed model of the IR sensor and a target facet model to generate realistic scenes of targets in clutter. This simulation allows the target-to-background contrast to be adjusted and permits the inclusion a variety of discrete clutter types. A flow diagram for the STIG simulation is shown in figure 2.

In addition to the simulated imagery, we also had available for analysis data from two CFTs. One CFT was conducted in Yuma, Arizona. For this test, the clutter was dominated by large discretess—ocotillo shrubs. The second CFT was conducted in Huntsville, Alabama. For this test, the clutter was more benign. Figure 3 shows some examples of simulated UCIR imagery under a variety of conditions.

Clearly, this imagery presents a very challenging ATA problem. We note that the images shown in figure 3 have been amplitude-scaled for viewing purposes, so the object detection problem is even more difficult than it appears from the images. In the next section, we introduce the idea of image restoration and then, in the

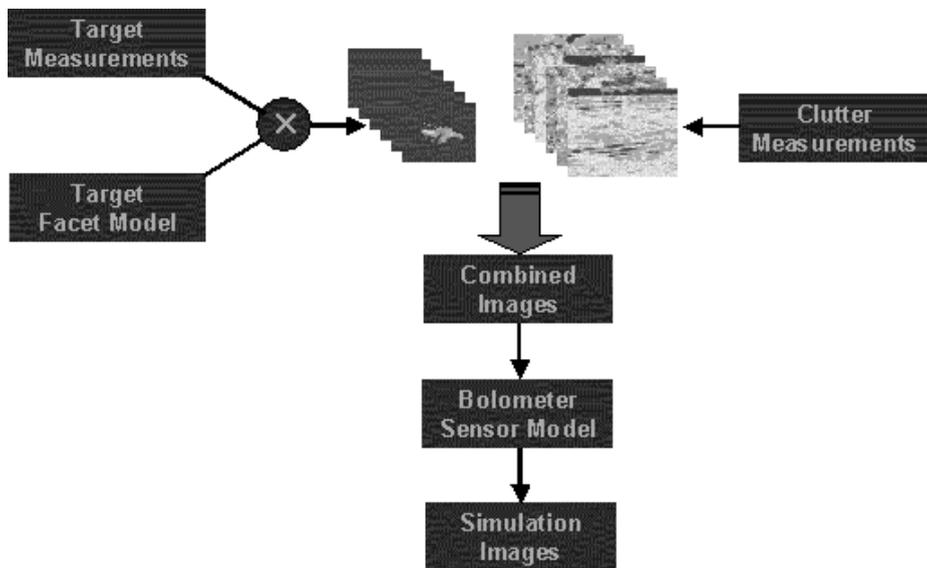


Figure 2. Simulated UCIR image generation.

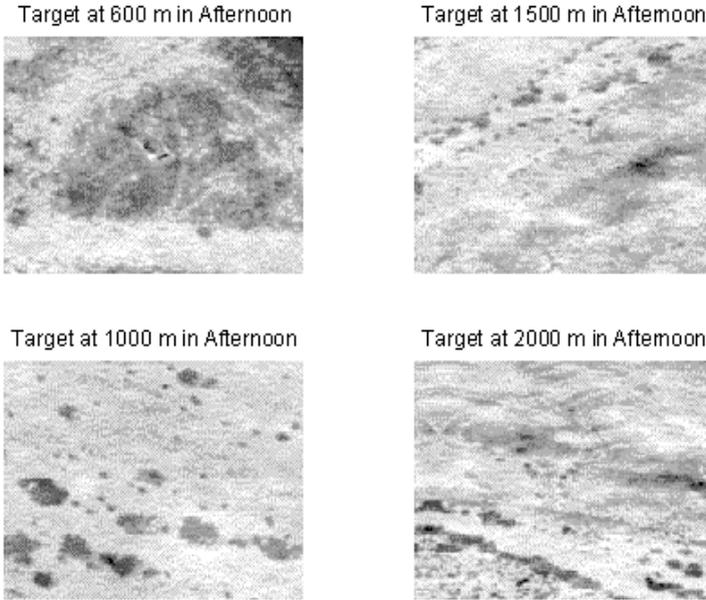


Figure 3. Sample UCIR imagery from simulation.

following section, we discuss one target detection algorithm that has shown some significant promise for this set of UCIR imagery.

4. Image restoration

Image restoration is a particular case of the more general field of inverse problems [4] that arise in many scientific fields [5]. In some cases, a direct approach of predicting the experimental data (i.e. the perceived image) using theoretical models from first principles may be applied. In the case of uncooled IR images, however, one of the fundamental characteristics of this inverse problem is that the theoretical understanding of the formation of these entities is either somewhat limited or, more often, not amenable to direct modelling with a great degree of accuracy.

The restoration problem can be cast in the form of a Fredholm integral equation of the first kind:

$$\mathbf{g}(x) = \int \mathbf{H}(x,y)\mathbf{f}(y)dy \tag{1}$$

In equation (1), $\mathbf{f}(y)$ is the function of interest, $\mathbf{g}(x)$ is the function accessible to measurement, and $\mathbf{H}(x,y)$ is the kernel of the integral equation. Equation (1) is the basic mathematical form of the restoration problem. A general form that may be used for the class of imaging systems where the spatial degradation can be modelled by a linear-shift-invariant impulse response is given as:

$$f_O(x,y) = \int_{-\infty}^{\infty} f_I(\alpha,\beta)h_D(x-\alpha,y-\beta)d\alpha d\beta + n(x,y) \tag{2a}$$

or

$$f_O(x, y) = f_I(x, y) * h_D(x, y) + n(x, y) \tag{2b}$$

In equation (2a), $f_O(x, t)$ is the observed image field, $f_I(\alpha, \beta)$ is an ideal image passed through a linear spatial degradation system with an impulse response $h_D(x - \alpha, y - \beta)$, and $n(x, y)$ is additive noise.

4.1. Examples of algorithms

The linear-shift-invariant filter defined by the impulse response $h_R(x, y)$ may be assumed as a model for the restoration system when the spatial degradation may be approximated by equation (2). In this case, the reconstructed image becomes:

$$\hat{f}_I(x, y) = \int_{-\infty}^{\infty} f_O(\alpha, \beta) h_R(x - \alpha, y - \beta) d\alpha d\beta \tag{3a}$$

or

$$f_I(x, y) = f_O(x, y) * h_R(x, y) \tag{3b}$$

After substituting and Fourier transforming, the continuous reconstruction may be represented as:

$$\hat{F}_I(\omega_x, \omega_y) = [F_I(\omega_x, \omega_y) H_D(\omega_x, \omega_y) + N(\omega_x, \omega_y)] H_R(\omega_x, \omega_y) \tag{4}$$

Examples of restoration algorithms include the simple inverse filter, the Wiener filter and constrained restoration. Constrained restoration is discussed in more detail in the next subsection.

4.2. Image restoration using constrained restoration

The image restoration techniques described above consider the images as arrays of numbers. A natural extension to these described is to impose physical constraints on the image produced by the restoration process. The physics of the image formation process constrain pixel amplitudes to be non-negative. Also, natural scenes are generally smooth. A discrete form will now be assumed for the degradation model. A vector-space model for this representation of image degradation

$$\mathbf{g} = \mathbf{D}\mathbf{f} + \mathbf{n} \tag{5}$$

where \mathbf{g} is a vector representation of the observed image, \mathbf{D} is a matrix representing the degradation function, \mathbf{f} is the ideal image and \mathbf{n} is additive noise.

The constrained restoration approach is to minimize the cost function

$$(\mathbf{g} - \mathbf{D}\hat{\mathbf{f}})^T \mathbf{M}(\mathbf{g} - \mathbf{D}\hat{\mathbf{f}}) \tag{6}$$

with the constraint

$$\hat{\mathbf{f}}^T \mathbf{S}\hat{\mathbf{f}} = d \tag{7}$$

where \mathbf{S} is a smoothing matrix, \mathbf{M} is an error weighting matrix and the scalar d is the degree of smoothing. The error weighting matrix is often chosen to be the inverse of the noise covariance matrix (K_N). The solution is obtained by applying the method of Lagrange multipliers. The Lagrangean function is:

$$\mathbf{G}(\hat{\mathbf{f}}, \lambda) = (\mathbf{g} - \mathbf{D}\hat{\mathbf{f}})\mathbf{M}^T(\mathbf{g} - \mathbf{D}\hat{\mathbf{f}}) + \lambda(\hat{\mathbf{f}}^T \mathbf{S}\hat{\mathbf{f}} - d) \tag{8}$$



Figure 4. Application of the constrained restoration filter.

After differentiating equation (6) with respect to f and λ and setting the derivatives to zero, the stationary point is found to be:

$$\hat{\mathbf{f}} = \mathbf{S}^{-1} \mathbf{D}^T (\mathbf{D} \mathbf{S}^{-1} \mathbf{D}^T + \lambda \mathbf{M}^{-1})^{-1} \mathbf{g} \quad (9)$$

The Lagrange multiplier in equation (8) can be chosen to create a satisfactory compromise between restoration error and smoothness. The results of applying the constrained restoration filter to an uncooled IR image degraded by out-of-focus optics and high levels of sensor noise are shown in figure 4. Note that the result obtained using the larger Lagrange multiplier is smoother.

4.3. Wavelet-based restoration

Although the Wiener filtering is the optimal trade-off of inverse filtering and noise smoothing, in the case when the spatial degradation filter is singular, the Wiener filter can actually amplify the noise. Wavelet-based denoising provides a technique for removing the amplified noise. The overall reconstruction technique employs Fourier-domain filtering and wavelet-domain regularization. The regularized inverse filter looks like a modified Wiener filter:

$$H_{\alpha}(\omega_x, \omega_y) = \frac{H_D^*(\omega_x, \omega_y) W_{F_I}(\omega_x, \omega_y)}{|H_D(\omega_x, \omega_y)|^2 W_{F_I}(\omega_x, \omega_y) + \alpha W_N(\omega_x, \omega_y)} \quad (10)$$

where α can be selected optimally to minimize the overall mean-square error. Since wavelet transforms have good decorrelation properties, the wavelet coefficients of the image can be formulated in a stochastic model, and the power spectrum can be better estimated. This implies that both inverse filtering and noise smoothing can be performed in the wavelet domain. Specifically, the power spectrum of the image in a same subband can be estimated under the assumption that the wavelet coefficients are independent. Therefore, the power spectrum is just the variance of the wavelet coefficients. This particular approach (i.e. regularization in the wavelet domain) is only valid under the important assumption that an undecimated wavelet transform is used and the spatial degradation function is separable. This regularized inverse filter/wavelet transform restoration is implemented in the form shown in figure 5.

If the above assumption is not valid then the implementation follows the form shown in figure 6.

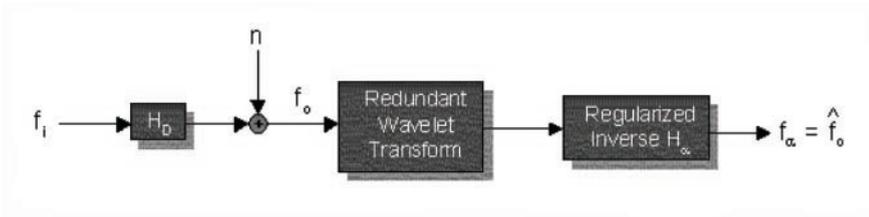


Figure 5. Wavelet-based restoration entirely in wavelet domain.

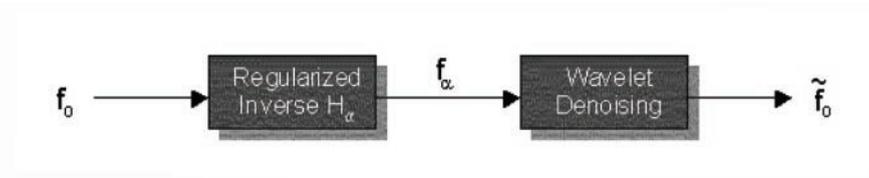


Figure 6. Generalized wavelet-based restoration.

5. LSVD algorithm description

In this section, we discuss the two primary algorithms used in processing the UCIR imagery. The LSVD algorithm was developed for background estimation and anomaly detection.¹ The LSVD algorithm is based on local statistical analyses and is a two-pass algorithm. The basic idea is to assign to each pixel an image block centred at the pixel. In this way, it is possible to construct a distance metric from one pixel (viewed as the distance between blocks) to another and, in particular, it is possible to introduce a distance from each pixel to a selected background region. Those pixels that are above a certain distance from the background are labeled as anomalies. Moreover, by selecting a background region, it is possible to estimate its local statistics and, in particular, to construct a set of singular values and right/left singular vectors. The right singular vectors can be used to ‘rotate’ each pixel in the image into the background coordinate system. The advantage of using this new coordinate system is that there are only a very few significant singular values for patterns that are man-made. We note that, since we are embedding a two-dimensional image in a high-dimensional pattern space, no more than two singular vectors are expected. The background clutter will have somewhat more complexity, thereby simplifying the task of isolating and identifying pixels that are significantly different from the local background. A flow diagram for the LSVD algorithm is given in figure 7.

Sample output of the LSVD algorithm is shown in figure 8.

6. Results

The STIG was used to generate a set of imagery with a clutter background collected from Yuma Proving Grounds. The data sets consisted of two different subsets from the morning and afternoon. For each of these cases, we generated

¹ Contact Fast Mathematical Algorithms and Hardware at <http://www.fmah.com>

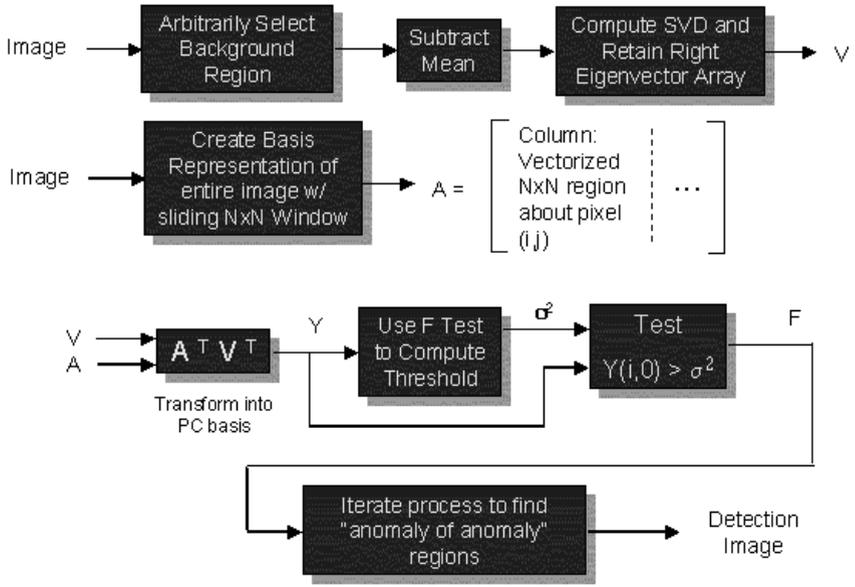


Figure 7. LSVD flow diagram.



Figure 8. (a) Original image; (b) first anomalies; (c) second anomalies.

images at two different acquisition ranges—1000m and 1500m—and with different aspect angles—0°, 150° and 300°. Within these image sets, the target-to-background contrast was selected randomly. Results are presented in figure 9. The simulated imagery used a 320 × 240, 25-micron pitch, VOx Bolometer array for the UCIR sensor model.

7. Conclusions

An excellent source of information regarding the current state-of-the-art for uncooled sensors can be found on the DARPA Microelectronics Technology Office web site [6]. While the LSVD algorithm has shown great promise as a detection algorithm for uncooled IR imagery, it is somewhat expensive computationally. This is a particularly significant drawback for many of the potential weapons where uncooled IR sensors have the most utility—small, smart munitions. In general, these weapons are severely constrained in volume, power, and

	R =1000m Morning	R =1000m Noon	R =1500m Morning	R =1500m Noon
# Images Processed	126	126	126	126
#Tgts Detected (Pd)	126 (1.0)	126 (1.0)	125 (.992)	109 (.865)
#False Alarms (FA/img)	11 (0.09)	571 (4.53)	165 (0.32)	1532 (12.16)
FAR (#FA/m ²)	1.53e-5	7.92e-4	1.02e-4	9.45e-4
# Images Processed	60	60	60	60
#Tgts Detected (Pd)	60 (1.0)	59 (.983)	60 (1.00)	50 (.833)
#False Alarms (FA/img)	2 (0.03)	452 (7.53)	7 (0.12)	740 (12.33)
FAR (#FA/m ²)	5.83e-6	1.32e-7	9.06e-6	9.58e-4

Figure 9. Processing results for simulated UCIR imagery.

cost. As a result, they have limited processing capabilities and specialized accelerators are not really an option. We are exploring a number of strategies to optimize the performance of LSVD. Since the LSVD algorithm does not use the singular values and requires only (a subset) of the right eigenvectors, we can replace the full SVD with a partial SVD [7]; however, the computational savings have not proven to be significant. A much more promising approach is to process the image in parallel. For example, the image can be broken down into subimages (with appropriate pixel overlap) and each subimage processed separately. As long as the subimages are large enough that the background can be estimated sufficiently well, there is very little degradation in the performance of the LSVD algorithm.

We have also been examining a number of other algorithms and approaches to improve the object detection performance. One particularly promising approach is a multiscale algorithm that uses complex (directional) wavelets. Results for this approach as a target detection algorithm have been presented elsewhere [8]. We are investigating the *Borrowed Strength Algorithm* (BSA) [9] as a post-detection algorithm for the UCIR imagery. A partial differential equation based approach to image processing, anisotropic diffusion or nonlinear diffusion filtering [10–16], is being evaluated for use on the UCIR imagery. We are also looking at the use of blind deconvolution or phase retrieval approaches [17], although they tend to suffer from numerical difficulties [18–20]. Longer term, we will investigate the incorporation of physical knowledge into image reconstruction algorithms as well as multiplexing approaches being developed at Duke University and Fast Mathematical Algorithms & Hardware.

Acknowledgments

This work was supported in part by the DARPA Defense Sciences Office under AFOSR contract F49620-98-C-0034. One of the authors (H.A.S.) would like to acknowledge useful discussions with F. Warner of FMAH, D. M. Healy of DAPRA and the University of Maryland, X.-B. Sun of Duke University, M. I. Miller of Johns Hopkins University, D. Casasent of Carnegie Mellon University, R. Haralick of the University of Washington and N. Intrator of Brown University.

We would also like to thank A. Morales of Raytheon Missiles Systems for providing the UCIR imagery sets.

References

- [1] TUMBAR, R., and BRADY, D. J., 2000, Southwest Symposium on Mixed-Signal Design, pp. 1–6.
- [2] DEVERSE, R. A., HAMMAKER, R. M., FATELEY, W. G., 2000, *Appl. Spectrosc.*, **54**, 1751.
- [3] SHANKAR, P. M., NEIFELD, M. A., and BRADY, D., 2001, OSA Annual Meeting and Exhibits 2001, Long Beach, CA, Octobr 14–18.
- [4] BERTERO, M., and BOCCACCI, P., 1998, *Introduction to Inverse Problems in Imaging* (Philadelphia, PA: IOP).
- [5] GROELSCH, C. W., 1993, *Inverse Problems in the Mathematical Sciences* (Weisbaden: Viewag).
- [6] <http://www.darpa.mil/not/uncooled/index.html>
- [7] HAIMI-COHEN, R., and COHEN, A., 1987, *IEEE Trans. Patt. Anal. Mach. Intell.*, **9**, 137.
- [8] JAING, Q., and KADAMBE, S., 2001, 35th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, November 4–7.
- [9] PRIEBE, C. E., MARCHETTE, D. J., and ROGERS, G. W., 1997, *IEEE Trans. Patt. Anal. Mach. Intell.*, **19**, 494.
- [10] WEICKERT, J., 1998, *Anisotropic Diffusion in Image Processing*.
- [11] COTTET, G.-H., and GERMAIN, L., 1993, *Math. Computat.*, **61**, 659.
- [12] WEICKERT, J., ZUIDERVELD, K. J., TER HAAR ROMENY, B. M., and NIESSEN, W. J., 1997, Proc. 1997 IEEE International Conference on Image Processing (ICIP-97, Santa Barbara, October 26–29), **3**, pp. 396–399.
- [13] BLACK, M. J., SAPIRO, G., and MARIMONT, D. H., 1998, *IEEE Trans. Image Process.*, **7**, 421.
- [14] WEICKERT, J., TER HAAR ROMENY, B. M., and VIERGEVER, M. A., 1998, *IEEE Trans. Image Process.*, **7**, 398.
- [15] WEICKERT, J., 2000, Technical Report 3/2000, Computer Science Series, February Computer Vision, Graphics, and Pattern Recognition Group, Dept. of Mathematics and Computer Science, University of Mannheim, Germany.
- [16] ZHU, S. C., and MUMFORD, D., 1998, Proc. International Conference on Computer Vision, Bombay, India, January.
- [17] FIENUP, J. R., 1991, in *International Trends in Optics*, edited by J. W. Goodman (Academic Press), pp. 407–422.
- [18] FIENUP, J. R., and WACKERMAN, C. C., 1986, *J. Opt. Soc. Am.*, **A3**, 1897.
- [19] FIENUP, J. R., 1984, in *Advances in Computer Vision and Image Processing, Vol. 1, Image Reconstruction from Incomplete Observations*, edited by T. S. Huang (Greenwich, CN: JAI Press), pp. 191–225.
- [20] FIENUP, J. R., 1982, *Appl. Opt.*, **21**, 2758.