Wavelet-Based Multi-Channel Image Denoising Using Fuzzy Logic

Jamal Saeedi and Ali Abedi

Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran
{jamal_saeidi,ali_abedi}@aut.ac.ir

Abstract. In this paper, we propose a new wavelet shrinkage algorithm based on fuzzy logic for multi-channel image denoising. In particular, intra-scale dependency within wavelet coefficients is modeled using a fuzzy feature. This feature space distinguishes between important coefficients, which belong to image discontinuity and noisy coefficients. Besides this fuzzy feature, we use inter-relation between different channels for improving the denoising performance compared to denoising each channel, separately. Then, we use the Takagi-Sugeno model based on two fuzzy features for shrinking wavelet coefficients. We examine our multi-channel image denoising algorithm in the dual-tree discrete wavelet transform domain, which is the new shiftable and modified version of discrete wavelet transform. Extensive comparisons with the state-of-the-art image denoising algorithms indicate that our image denoising algorithm has a better performance in noise suppression and edge preservation.

Keywords: Dual-tree discrete wavelet transform; Fuzzy membership function; Multi-channel image.

1 Introduction

Denoising has become an essential step in image analysis. Indeed, due to sensor imperfections, transmission channels defects, as well as physical constraints, noise weakens the quality of almost every acquired image.

Reducing the noise level is the main goal of an image denoising algorithm, while preserving the image features (such as edges, textures, etc.). The multi-resolution analysis performed by the wavelet transform has been shown to be a powerful tool to achieve these goals [1]. Indeed, in the wavelet domain, the noise is uniformly spread throughout the coefficients, while most of the image information is concentrated in the few largest ones. Classical wavelet-based denoising methods consist of three steps: (a) Compute the discrete wavelet transform (DWT), (b) Remove noise from the wavelet coefficients, and (c) Reconstruct the enhanced image by using the inverse DWT.

In this paper, we consider image with C channels. Typically, C is equal to three channels for RGB images, but for biological (fluorescence) and multi-channel satellite images, C might be larger. We denote these multi-channel images by:

$$X = [x(i, j,1),\ x(i, j,2) \ldots\ x(i, j, C)] \tag{1}$$

A. Elmoataz et al. (Eds.): ICISP 2010, LNCS 6134, pp. 44–53, 2010.

These images are corrupted by an additive Gaussian white noise following a normal law defined by a zero mean and a known σ^2 variance, that is $n \sim N(0, \sigma^2)$:

$$N = [n(i, j,1), n(i, j,2) \ldots n(i, j, C)] \tag{2}$$

We denote the resulting noisy image by:

$$Y = X + N \tag{3}$$

Due to the linearity of the wavelet transform, additive noise in the image domain remains additive in the transform domain. If $y_{s,d}(i, j)$ and $x_{s,d}(i, j)$ denote the noisy and the noise-free wavelet coefficients of scale s and orientation d, respectively, then we can model the additive noise in the transform domain as:

$$y_{s,d}(i, j, c) = x_{s.d}(i, j, c) + n_{s,d}(i, j, c) \tag{4}$$

The easiest method for denoising multi-channel image is simply to apply an existing denoising algorithm separately in each channel. However, this solution is beyond optimal, due to the presence of strong common information between the various channels.

In literatures, inter-channel information in both spatial and wavelet domain is used for multi-channel image denoising. S. Schulte et al used a fuzzy filter in spatial domain for color image denoising. The filter consists of two sub-filters. The first sub-filter computes the fuzzy distance between the color components of the central pixel and its neighborhood. These distances determine to what degree each component must be corrected. The second sub-filter corrects pixels where the color components differences are corrupted so much that they appear as outliers in comparison to their environment [2]. Luisier's et al. multi-channel Sure-Let work follows their SURE-LET approach [3], where the denoising algorithm is parameterized as a linear expansion of thresholds (LET) and optimized using Stein's unbiased risk estimate (SURE). The proposed wavelet thresholding function is point-wise and depends on the coefficients of same location in the other channels, as well as on their parents in the coarser wavelet sub-band [4]. Pizurica et al. have also adapted their original ProbShrink by including the inter-channel dependencies in the definition of their local spatial activity indicator [5].

In image denoising, where a trade-off between noise suppression and the maintenance of actual image discontinuity must be made, solutions are required to detect important image details and accordingly adapt the degree of noise smoothing. We always postulate that noise is uncorrelated in the spatial domain; it is also uncorrelated in the wavelet domain. With respect to this principle, we use a fuzzy feature to enhance image information in wavelet sub-bands. This feature space distinguishes between important coefficients, which belong to image discontinuity and noisy coefficients. Besides this fuzzy feature, we use inter-relation between different channels for improving the denoising performance compared to denoising each channel, separately. Then, we use the Takagi-Sugeno model [6] based on two fuzzy features for shrinking wavelet coefficients.

In addition, we examine our image denoising algorithm in the dual-tree DWT (DT-DWT), which provides both shiftable sub-bands and good directional selectivity and low redundancy [7].

This paper is structured as follows: Section 2 presents proposed multi-channel image denoising algorithm. Section 3 gives results and the comparisons. Finally, we conclude with a brief summary in Section 4.

2 Multi-Channel Image Denoising

In this Section, we describe our multi-channel image denoising algorithm. First, the input noisy image is decomposed using DT-DWT. After extracting two fuzzy features from different sub-bands at different scales, wavelet coefficients are shrunk using a simple fuzzy rule. Finally, inverse DT-DWT of modified wavelet coefficients generates denoised image.

2.1 Single and Multi-Channel Fuzzy Features

In multi-channel images, different channels are correlated: an image discontinuity from one channel is probably to occur in at least some of remaining channels. As for this reason and uncorrelated nature of noise, we use a new feature for denoising of multi-channel images. Improving wavelet coefficients' information in the shrinkage step is the main idea.

The first feature is obtained using a nonlinear averaging filter in the wavelet sub-bands of each single channel. We want to give large weights to neighboring coefficients with similar magnitude, and vice versa. The larger coefficients, which produced by noise are always isolated or unconnected, but edge coefficients are clustered and persistence. It is well known that the adjacent points are more similar in magnitude. Therefore, we use a fuzzy function $m(l,k,c)$ of magnitude similarity and a fuzzy function $s(l,k,c)$ of spatial similarity, which is defined as:

$$m(l,k,c) = \exp\left(-\left(\frac{y_{s,d}(i,j,c) - y_{s,d}(i+l,j+k,c)}{Thr}\right)^2\right) \tag{5}$$

$$s(l,k,c) = \exp\left(-\left(\frac{l^2+k^2}{N}\right)\right) \tag{6}$$

where $y_{s,d}(i,j,c)$ and $y_{s,d}(i+l,j+k,c)$ are central coefficient and neighbor coefficients in wavelet sub-bands of channel c, respectively. $Thr = K \times \hat{\sigma}_n^c, 3 \le K \le 4$, $\hat{\sigma}_n^c$ is estimated noise variance of channel c using median estimator, and N is the number of coefficients in the local window $k \in [-K...K]$, and $l \in [-L...L]$.

According the two fuzzy functions, we can get adaptive weight $w(l,k,c)$ for each neighboring coefficient in each channel:

$$w(l,k,c) = m(l,k,c) \times s(l,k,c) \tag{7}$$

Using the adaptive weights $w(l,k,c)$, we can obtain the first fuzzy feature for each coefficient in wavelet sub-bands as follows:

$$f_1(i,j,c) = \frac{\sum_{l=-L}^{L}\sum_{k=-K}^{K} w(l,k,c)\times\left|y_{s,d}(i+l,j+k,c)\right|}{\sum_{l=-L}^{L}\sum_{k=-K}^{K} w(l,k,c)} \tag{8}$$

In addition, the second fuzzy feature that we have used for multi-channel image denoising is defined as:

$$f_2(i,j) = \frac{\sum_{c=1}^{C}\left(\sum_{l=-L}^{L}\sum_{k=-K}^{K}\left(\prod_{c=1}^{C} w(l,k,c)\right)\times\left|y_{s,d}(i+l,j+k,c)\right|\right)}{C\times\sum_{l=-L}^{L}\sum_{k=-K}^{K}\left(\prod_{c=1}^{C} w(l,k,c)\right)} \tag{9}$$

The second fuzzy feature equally is used to shrink wavelet coefficients in each channel. In fact, for denoising one channel, the second feature simultaneously takes the remaining channels' information into account in the denoising process. Then, we use the Takagi-Sugeno model based on the two fuzzy features for shrinking wavelet coefficients [7].

2.2 Fuzzy Shrinkage Rule

Having the fuzzy features, we will define Linguistic IF-THEN rules for shrinking wavelet coefficients in each channel as follows:

IF $f_1(i,j,c)$ (the first fuzzy feature) is large **AND** $f_2(i,j)$ (the second fuzzy feature) is large **THEN** shrinkage of wavelet coefficient in channel c i.e. $y_{s,d}(i,j,c)$ is small.

The idea behind this simple fuzzy rule is to assign small shrinkage weights to the wavelet coefficients, which have similar values to the neighbors and corresponding coefficients in the remaining channels. For color images, it is important to treat pixels as color components not as three separate colors. When only the separate channels are considered, more artifacts are introduced.

For building standard rule from linguistic rule, each linguistic value is represented by a membership function. The membership functions that are used to represent the two fuzzy sets of large $f_1(i,j,c)$ (the first fuzzy feature) and large $f_2(i,j)$ (the second fuzzy feature), are denoted as $\mu(f_1(i,j,c))$ and $\mu(f_2(i,j))$, respectively.

Here, we use the spline-based curve as fuzzy membership function, which is a mapping on the vector x, and is named because of its S-shape. The parameters T_1 and T_2 locate the extremes of the sloped portion of the curve, as given by:

$$\mu(x) = \begin{cases} 0 & x \le T_1 \\ 2 \times \left(\dfrac{x - T_1}{T_2 - T_1} \right)^2 & T_1 < x \le \dfrac{T_1 + T_2}{2} \\ 1 - 2 \times \left(T_2 - \dfrac{x}{T_2 - T_1} \right)^2 & \dfrac{T_1 + T_2}{2} \le x < T_2 \\ 1 & x \ge T_2 \end{cases} \tag{10}$$

For the pair of values $\{\mu(f_1(i,j,c)), \mu(f_2(i,j))\}$, the degree of satisfaction of the antecedent part of the rule determines the firing strength of the rule:

$$\tau(i,j,c) = f_1(i,j,c) \text{ AND } f_2(i,j) \tag{11}$$

The AND operation is typically implemented as *minimum*, but any other triangular norms (t-norm) may be used (such as *algebraic product*, *weak*, and *bounded sum*). Originally, t-norms appeared in the context of probabilistic metric spaces [8]. Then they were used as a natural interpretation of the conjunction in the semantics of mathematical fuzzy logics and they are used to combine criteria in multi-criteria decision making [9]. In such applications, the *minimum* or *product* t-norm are usually used because of a lack of motivation for other t-norms [8]. We have chosen *algebraic product* for the AND operation, because it is the standard semantics for strong conjunction in product fuzzy logic and is continues as a function. Finally, the estimated noise free signal is obtained using the following formula:

$$\hat{x}_{s,d}(i,j,c) = \tau(i,j,c) \times y_{s,d}(i,j,c) \tag{12}$$

where s and d are scale and orientation of DT-DWT sub-bands, and $\tau(i,j,c)$ is obtained using (11).

In addition, for building fuzzy membership function we used test images in different noise levels, to obtain best values for the thresholds.

2.3 Post-processing

Processing artifact usually result from a modification of the spatial correlation between wavelet coefficients (often caused by zeroing of small neighboring coefficients) or using DWT, which is shift invariance and will cause some visual artifacts (such as Gibbs phenomena) in thresholding-based denoising. For this reason, we use a fuzzy filter on the results of our fuzzy-shrink algorithm to reduce artifacts to some extent. First, we use a window of size $(2L+1)\times(2K+1)$ centered at (i,j) to filter the current image pixel at position (i,j). Next, we calculate the similarity of neighboring pixels to the center pixel using following formula:

$$m(l,k) = \prod_{c=1}^{C} \exp\left(-\left(\frac{\hat{X}(i,j,c) - \hat{X}(i+l, j+k, c)}{Thr} \right)^2 \right) \tag{13}$$

where $\hat{X}(i,j,c)$ and $\hat{X}(i+l, j+k, c)$ are central pixel and neighbor pixels in the denoised image of channel c, respectively. N is the number of pixels in the local window $k \in [-K...K]$, and $l \in [-L...L]$, and $2.55 < Thr < 7.65$.

According the fuzzy feature, we can get adaptive weight $w(l,k)$ for each neighboring coefficient:

$$w(l,k) = m(l,k) \times s(l,k) \tag{14}$$

where $s(l,k)$ is obtained using (6).

Finally, the output of post-processing step is determined as follows:

$$\tilde{X}(i,j,c) = \frac{\sum_{l=-L}^{L} \sum_{k=-K}^{K} w(l,k) \times \hat{X}(i+l, j+k, c)}{\sum_{l=-L}^{L} \sum_{k=-K}^{K} w(l,k)} \tag{15}$$

where $\hat{X}$ is the denoised image, which is obtained via our fuzzy-shrink algorithm.

3 Experimental Results

In the satellite systems, it may be desirable to perform denoising before the image compression step in order to improve the compression efficiency. Therefore, we use both color and multi-spectral satellite images for experimental dataset, which was captured by simulated additive Gaussian white noise.

In this section, we compare our multi-channel fuzzy denoising method with three state-of-the-art wavelet-based techniques: Pizurica's multi-channel ProbShrink [5] and Luisier's et al. multi-channel Sure-Let [4]. In addition, we compare our multi-channel image denoising algorithm with the Portilla's BLS-GSM [10], which denoised each channel, separately.

It should be mentioned that for comparing ProbShrink [5] and Sure-Let [4] methods, which are proposed in the discrete wavelet transform domain, we use a critically sampled orthonormal wavelet basis with eight vanishing moments (sym8) over four decomposition stages in our fuzzy-shrink method. On the other hand, BLS-GSM method has used a redundant wavelet transform (an 8-orientations full steerable pyramid [10]). Therefore, for a fair comparison between BLS-GSM and our fuzzy-shrink methods, we use dual-tree discreet wavelet transform over four decomposition stages for wavelet analysis. We measured the experimental results by the peak signal-to-noise ratio (PSNR) in decibels (dB), objectively:

$$PSNR = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) \tag{16}$$

where

$$MSE = \frac{1}{C \times M \times N} \sum_{c=1}^{C} \left(\sum_{i=1}^{M} \sum_{j=1}^{N} \left(\tilde{X}(i,j,c) - X(i,j,c)\right) \right)^2 \tag{17}$$

where X is the original image, $\tilde{X}$ is the estimated noise free signal, and $M \times N$ is the number of pixels.

Table 1. Comparison of multi-channel image denoising algorithms

σ_n	5	15	25	30	50	5	15	25	30	50
Method	**Lena 512×512**					**Peppers 512×512**				
Input PSNR	34.15	24.69	20.37	18.86	14.82	34.25	24.78	20.47	18.98	15.01
Prob-Shrink-M (**DWT**)	35.61	31.40	29.31	28.51	26.08	34.86	30.35	28.12	27.82	25.29
Sure-let-M (**DWT**)	**37.86**	32.40	29.60	28.64	26.20	**36.62**	**31.79**	29.16	28.28	25.61
Fuzzy-Shrink-M (**DWT**)	36.81	32.08	30.01	29.31	26.77	36.13	31.26	29.17	28.49	25.86
	36.75	**32.65**	**30.60**	**29.80**	**27.14**	36.05	31.53	**29.61**	**28.95**	**26.15**
BLS-GSM **Best Redundant**	37.11	32.72	30.55	29.95	27.89	36.12	31.63	29.86	29.03	26.98
Fuzzy-Shrink-M (**DT-DWT**)	**37.56**	33.16	31.17	30.35	27.76	36.57	31.94	30.01	29.21	26.64
	37.42	**33.29**	**31.28**	**30.46**	**27.92**	36.41	**32.23**	**30.24**	**29.45**	**26.78**
Method	**Baboon 512×512**					**Southern California 512×512**				
Input PSNR	34.15	24.66	20.33	18.79	14.75	34.26	24.81	20.44	18.92	14.88
ProbShrink-M (**DWT**)	33.11	26.44	23.73	22.89	20.94	33.29	27.13	24.31	23.43	21.06
Sure-let-M (**DWT**)	**34.46**	26.82	24.03	23.27	21.43	33.36	27.74	24.95	24.05	21.66
Fuzzy-Shrink-M (**DWT**)	34.01	27.25	24.41	23.47	21.49	35.21	28.25	25.31	24.41	21.99
	33.86	**27.44**	**24.86**	**23.96**	**21.79**	**35.35**	**28.41**	**25.64**	**24.76**	**22.42**
BLS-GSM **Best Redundant**	**35.01**	27.56	24.75	24.01	21.62	35.31	28.31	25.32	24.68	22.13
Fuzzy-Shrink-M (**DT-DWT**)	34.33	**27.66**	25.25	24.57	22.28	**35.23**	28.94	26.15	25.23	22.68
	34.21	27.56	**25.31**	**24.72**	**22.45**	35.19	**29.21**	**26.41**	**25.48**	**22.84**

Note: Output PSNRs have been averaged over five noise realizations. The best redundant results are obtained using the BLS-GSM 3× 3 with an 8-orientations full steerable pyramid [10]. In addition, the results in the second row for our fuzzy method are obtained after post processing.

Table 1 summarizes the obtained results. As it can be observed in Table 1, our fuzzy multi-channel image denoising method is already competitive with the best techniques. When looking closer at the results, we observe the following:

- Our method (Fuzzy-Shrink using DWT) provides better results than Pizurica's multi-channel ProbShrink, which integrates the intra-scale dependencies (average gain of +1.199 dB).
- Our method (Fuzzy-Shrink using DWT) outperforms the Luisier's et al. multi-channel Sure-Let, which is used the inter-scale and inter-channel dependencies more than +0.449 dB on average.
- Our method (Fuzzy-Shrink using DT-DWT) also gives better results than Portilla's BLS-GSM (Best-Redundant), which is used for each channel separately (average gain of +0.435 dB).
- The worst result is obtained for the "Baboon" image. Our explanation for this is that the "Baboon" image is very noisy (i.e. texture area in this image is similar to noise), and when we use the fuzzy features for taking into account the neighbor dependency, it will be smoother in the resulting image.

In addition, Figs. 1 and 2 illustrate the results of noisy multispectral "southern California" and noisy color "Lena" images from different methods. As it can be observed in Figs. 1 and 2, our multi-channel image denoising is successful in preserving image edges and fewer artifacts as visual criteria as compared to other methods.

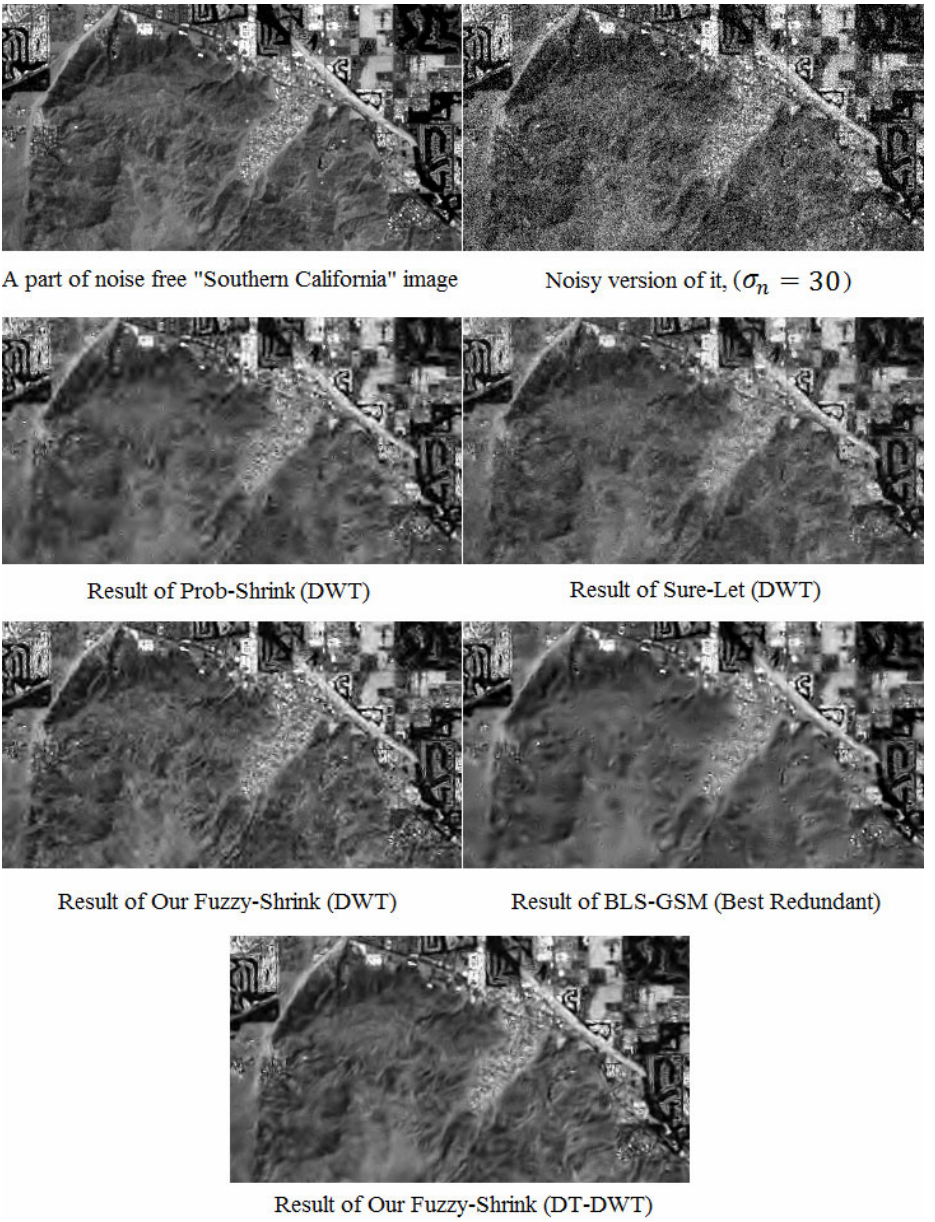

Fig. 1. Comparison of multi-channel image denoising algorithms

Fig. 2. Comparison of multi-channel image denoising algorithms

4 Conclusion

In this paper, we propose a new wavelet-based multi-channel image denoising using intra-scale dependency as a fuzzy feature, and inter-channel relation to improve wavelet coefficients' information at the shrinkage step. We use the DT-DWT for wavelet analysis, because it is shift invariant, and has more directional sub-bands

compared to the DWT. In other words, proposing a new method for shrinking wavelet coefficients in the second step of the wavelet-based image denoising is the main novelty of this paper.

The comparison of the denoising results obtained with our algorithm, and with the best state-of-the-art methods, demonstrate the performance of our fuzzy approach, which gave the best output PSNRs for most of the images. In addition, the visual quality of our denoised images exhibits the fewest number of artifacts and preserves most of edges compared to other methods.

References

1. Mallat, S.: A theory for multiresolution signal decomposition: The wavelet representation. IEEE Trans. Pattern Anal. Mach. Intell. 11(7), 674–693 (1989)
2. Schulte, S., Witte, V.D., Kerre, E.E.: A Fuzzy Noise Reduction Method for Color Images. IEEE Trans. Image Process 16(5), 1425–1436 (2007)
3. Blu, T., Luisier, F.: The SURE-LET approach to image denoising. IEEE Trans. Image Process 16, 2778–2786 (2007)
4. Luisier, F., Blu, T.: SURE-LET Multichannel Image Denoising: Interscale Orthonormal Wavelet Thresholding. IEEE Trans. Image Process 17(4), 482–492 (2008)
5. Pizurica, A., Philips, W.: Estimating the probability of the presence of a signal of interest in multiresolution single-and multiband image denoising. IEEE Trans. Image Process, 654–665 (2006)
6. Kingsbury, N.G.: Complex Wavelets for Shift Invariant Analysis and Signals. Applied and Computational Harmonic Analysis 10(3), 234–253 (2001)
7. Zadeh, L.A.: Fuzzy logic and its application to approximate reasoning. Inf. Process 74, 591–594 (1973)
8. Driankov, D., Hellendoorn, H., Reinfrank, M.: An Introduction to Fuzzy Control. Springer, Heidelberg (1993)
9. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer, Dordrecht (1998)
10. Portilla, J., Strela, V., Wainwright, M., Simoncelli, E.: Image denoising using Gaussian scale mixtures in the wavelet domain. IEEE Transactions on Image Processing, 1338–1351 (2003)